Software Defined Services (SDS) For High Performance Large Scale Science Data Streams Across 100 Gbps WANs

Joe Mambretti, Director, (j-mambretti@northwestern.edu)
International Center for Advanced Internet Research (www.icair.org)
Northwestern University
Director, Metropolitan Research and Education Network (www.mren.org)
Director, StarLight, PI StarLight SDX,Co-PI Chameleon, PI-iGENI, PI-OMNINet (www.startap.net/starlight)

The International Conference for High Performance Computing, Networking, Storage and Analysis (SC16)
Salt Lake City, Utah
November 13-17, 2016

iCAIR

STARLIGHT℠

GriPhyN
Data Intensive Science

NEPTUNE
A Fiber-Optic Telescope to Inner Space

NEES

GENOMES to LIFE
BIOLOGICAL SOLUTIONS FOR ENERGY CHALLENGES
U.S. DEPARTMENT OF ENERGY

NG Digital Sky Survey

NEON
National Ecological Observatory Network

ATLAS

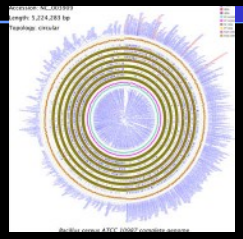
ALMA

ALMA: Atacama Large Millimeter Array

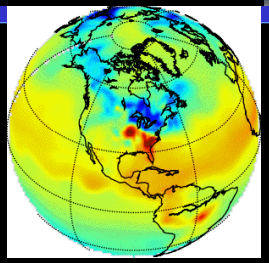
ANDRILL: Antarctic Geological Drilling
www.andrill.org

BIRN

BIRN: Biomedical Informatics Research Network
www.nbirn.net



CAMERA metagenomics
camera.calit2.net



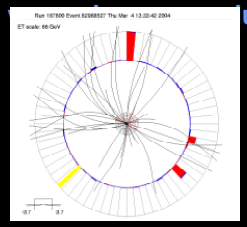
Carbon Tracker
www.esrl.noaa.gov/gmd/ccgg/carbontracker

CineGrid™

CineGrid
www.cinegrid.org

LHC ONE

LHCONE
www.lhcone.net



DØ (DZero)
www-d0.fnal.gov

GEON

GEON: Geosciences Network
www.geongrid.org

GLEON: Global Lake Ecological Observatory Network

OOI
OCEAN OBSERVATORIES INITIATIVE
CYBERINFRASTRUCTURE
Providing a link between ocean research and discovery

OOI-CI
ci.oceanobservatories.org

ISS: International Space Station
www.nasa.gov/station

CLASS

Comprehensive Large-Array Stewardship System
www.class.noaa.gov

LIGO
www.ligo.org

PRAGMA

Pacific Rim Applications and Grid Middleware Assembly
www.pragma-grid.net

WLCG
Worldwide LHC Computing Grid

WLCG
lcg.web.cern.ch/LCG/public/

TeraGrid™

TeraGrid
www.teragrid.org

IVOA

IVOA: International Virtual Observatory
www.ivoa.net

Open Science Grid

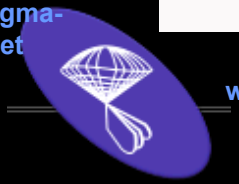
OSG
www.opensciencegrid.org

the globus alliance

Globus Alliance
www.globus.org

SKA
SQUARE KILOMETRE ARRAY

SKA
www.skatelescope.org



Sloan Digital Sky Survey
www.sdss.org

XSEDE

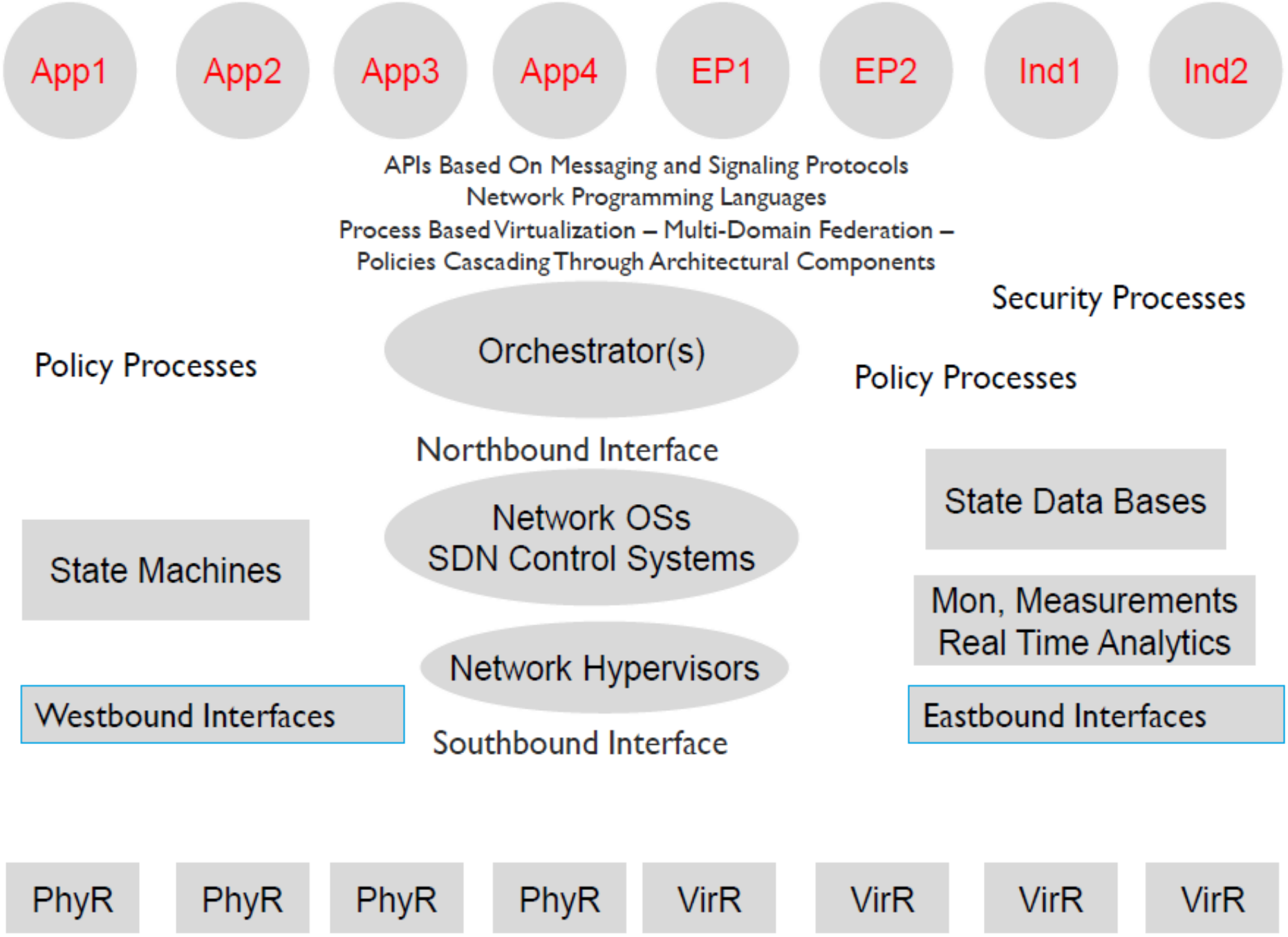
XSEDE
www.xsede.org

Compilation By Maxine Brown

STARLIGHT℠

Macro Network Science Themes

- Transition From Legacy Networks To Networks That Take Full Advantage of IT Architecture and Technology
- Extremely Large Capacity (Multi-Tbps Streams)
- High Degrees of Communication Services Customization
- Highly Programmable Networks
- Network Facilities As Enabling Platforms for Any Type of Service
- Network Virtualization
- Highly Distributed Processes


App1
App2
App3
App4
EP1
EP2
Ind1
Ind2
APIs Based On Messaging and Signaling Protocols
Network Programming Languages
Process Based Virtualization – Multi-Domain Federation –
Policies Cascading Through Architectural Components
Security Processes
Orchestrator(s)
Policy Processes
Policy Processes
Northbound Interface
State Data Bases
Network OSs
SDN Control Systems
State Machines
Mon, Measurements
Real Time Analytics
Network Hypervisors
Westbound Interfaces
Eastbound Interfaces
Southbound Interface
PhyR
PhyR
PhyR
PhyR
VirR
VirR
VirR
VirR

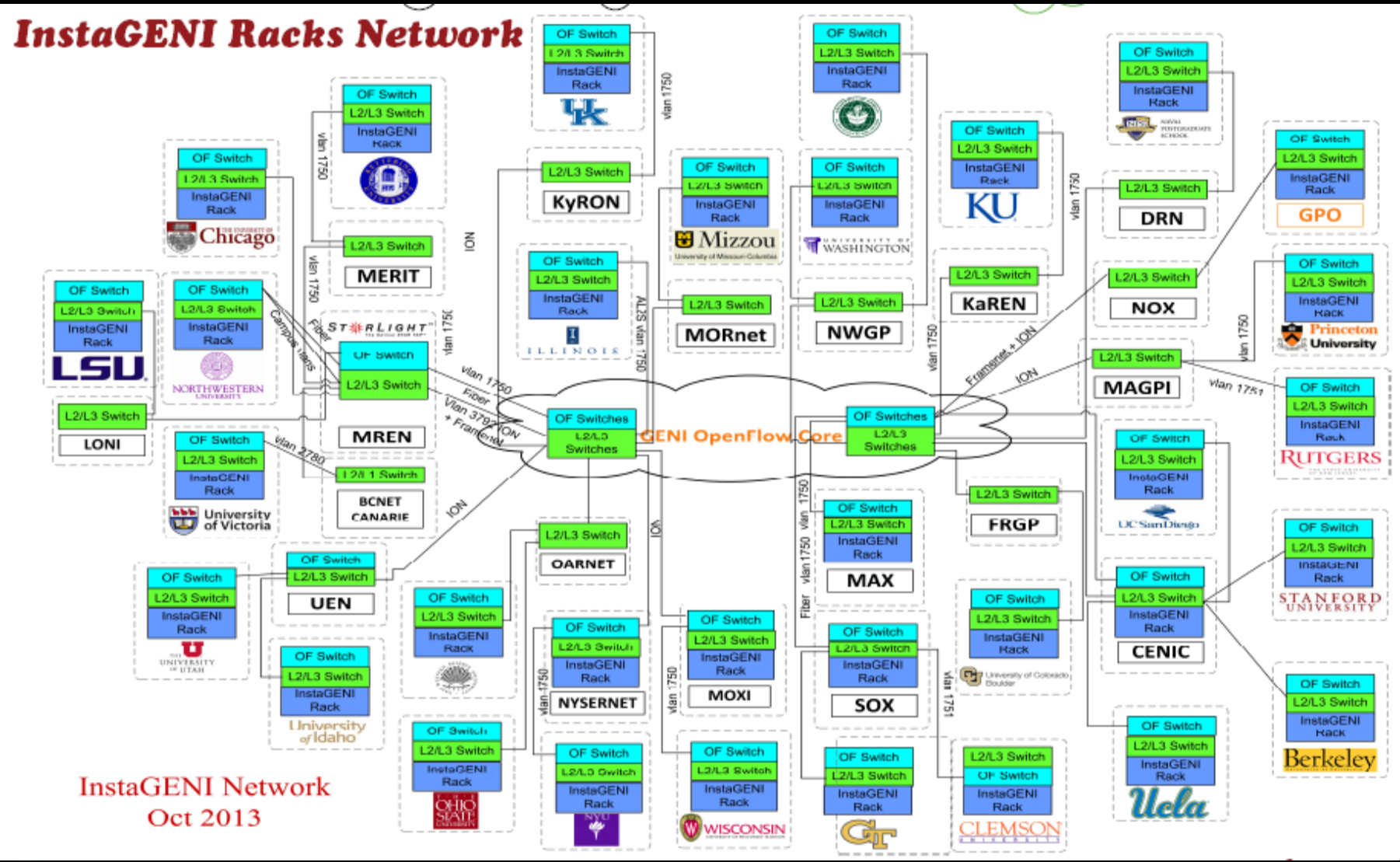
National Science Foundation's Global Environment for Network Innovations (GENI)

- GENI Is Funded By The National Science Foundation's Directorate for Computer and Information Science and Engineering (CISE)
- GENI Is a Virtual Laboratory For Exploring Future Internets At Scale.
- GENI Is Similar To Instruments Used By Other Science Disciplines, e.g., Astronomers – Telescopes, HEP - Synchrotrons
- GENI Creates Major Opportunities To Understand, Innovate and Transform Global Networks and Their Interactions with Society.
- GENI Is Dynamic and Adaptive.
- GENI Opens Up New Areas of Research at the Frontiers of Network Science and Engineering, and Increases the Opportunity for Significant Socio-Economic Impact.

Future Cyberinfrastructure

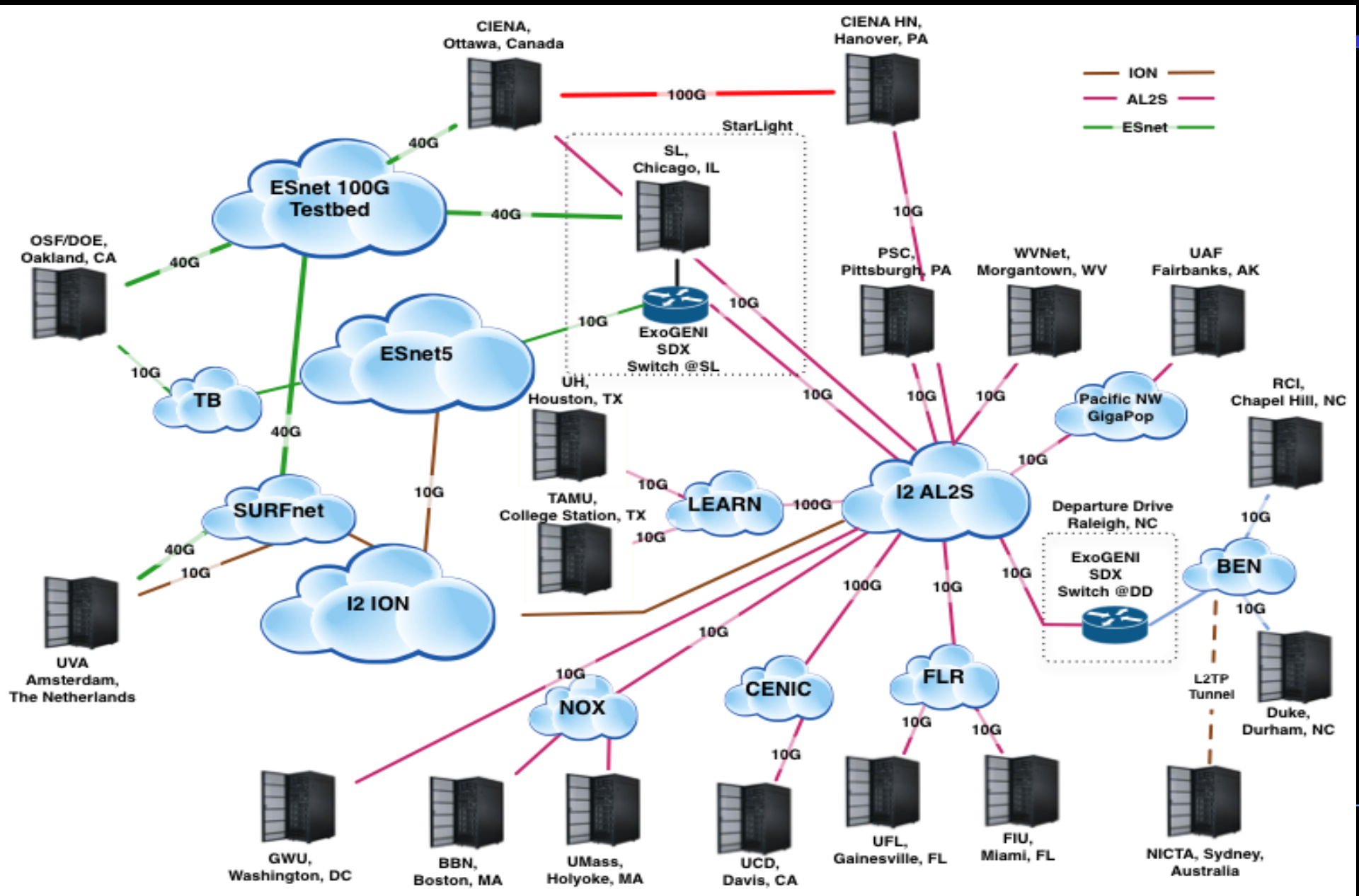
- Large Scale Highly Distributed Infrastructure That Can Support Multiple Empirical Research Testbeds At Scale
- Next Generation GENI, Edge Clouds, IOT, US Ignite, Platform for Advanced Wireless Research (PAWR) and Many Others
- Currently Being Planned – Will Be Designed, Implemented and Operated By Researchers for Researchers

National Science Foundation Global Environment for Network innovations



STARLIGHT SDX

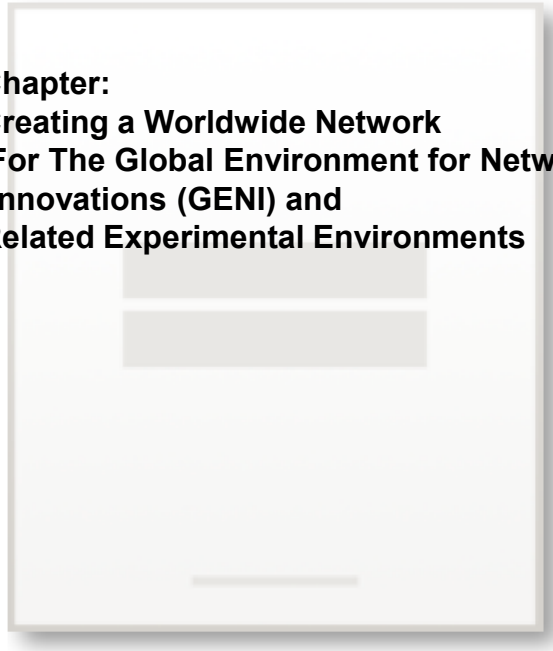
International 40G and 100 G ExoGENI Testbed



Springer

springer.com

Chapter:
Creating a Worldwide Network For The Global Environment for Network Innovations (GENI) and Related Experimental Environments



1st ed. 2016, XVIII, 655 p. 216 illus., 183 illus. in color.

Printed book

R. **McGeer**, M. **Berman**, C. **Elliott**, R. **Ricci** (Eds.)

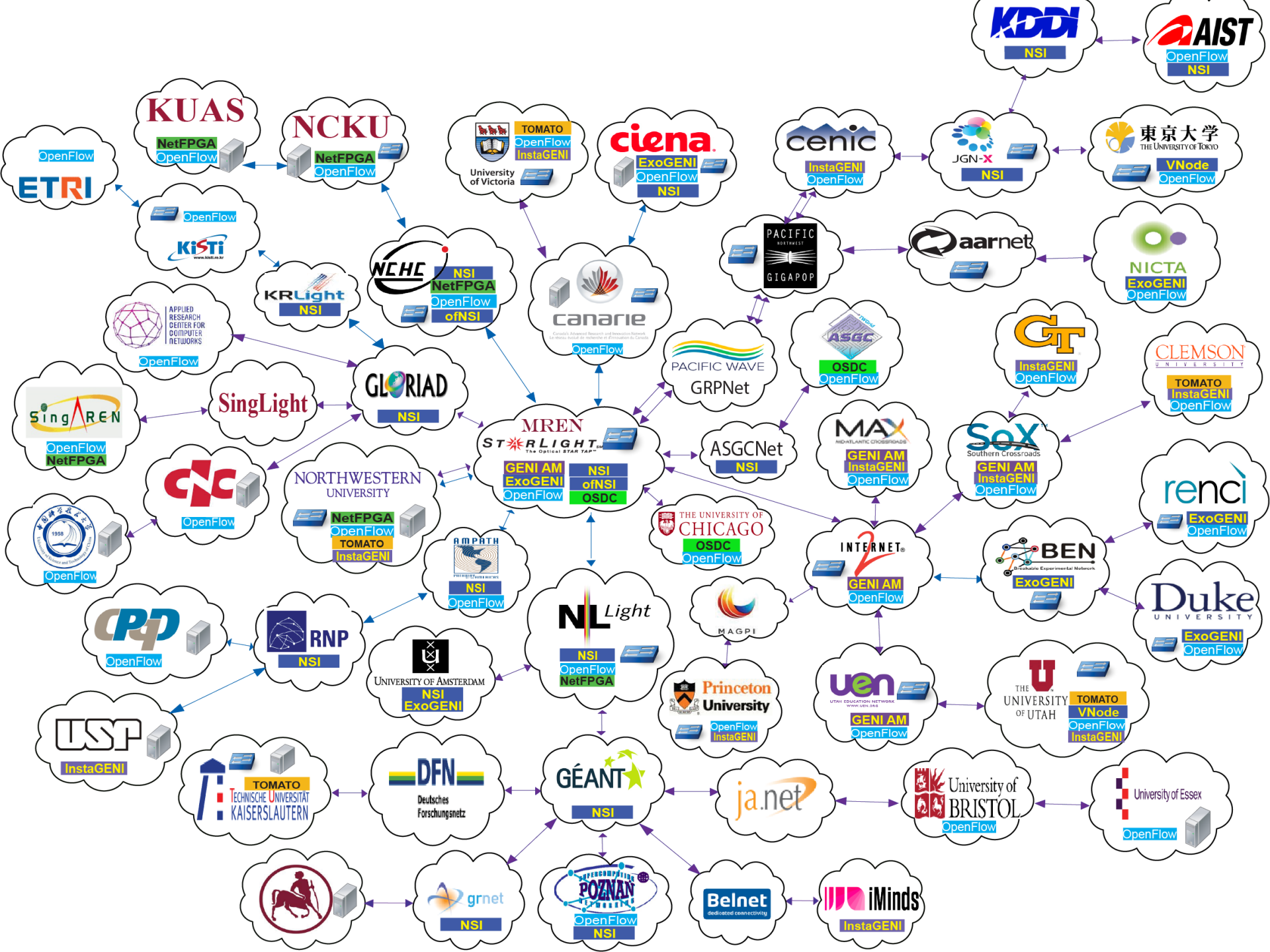
The GENI Book

- Provides a foundational overview of GENI's core architectural concepts
- Presents a detailed discussion of architecture and implementation
- Includes 24 chapters, divided into five sections, which outline GENI from precursors to architecture, development, applications, and then world federation
- Offers an extensive bibliography

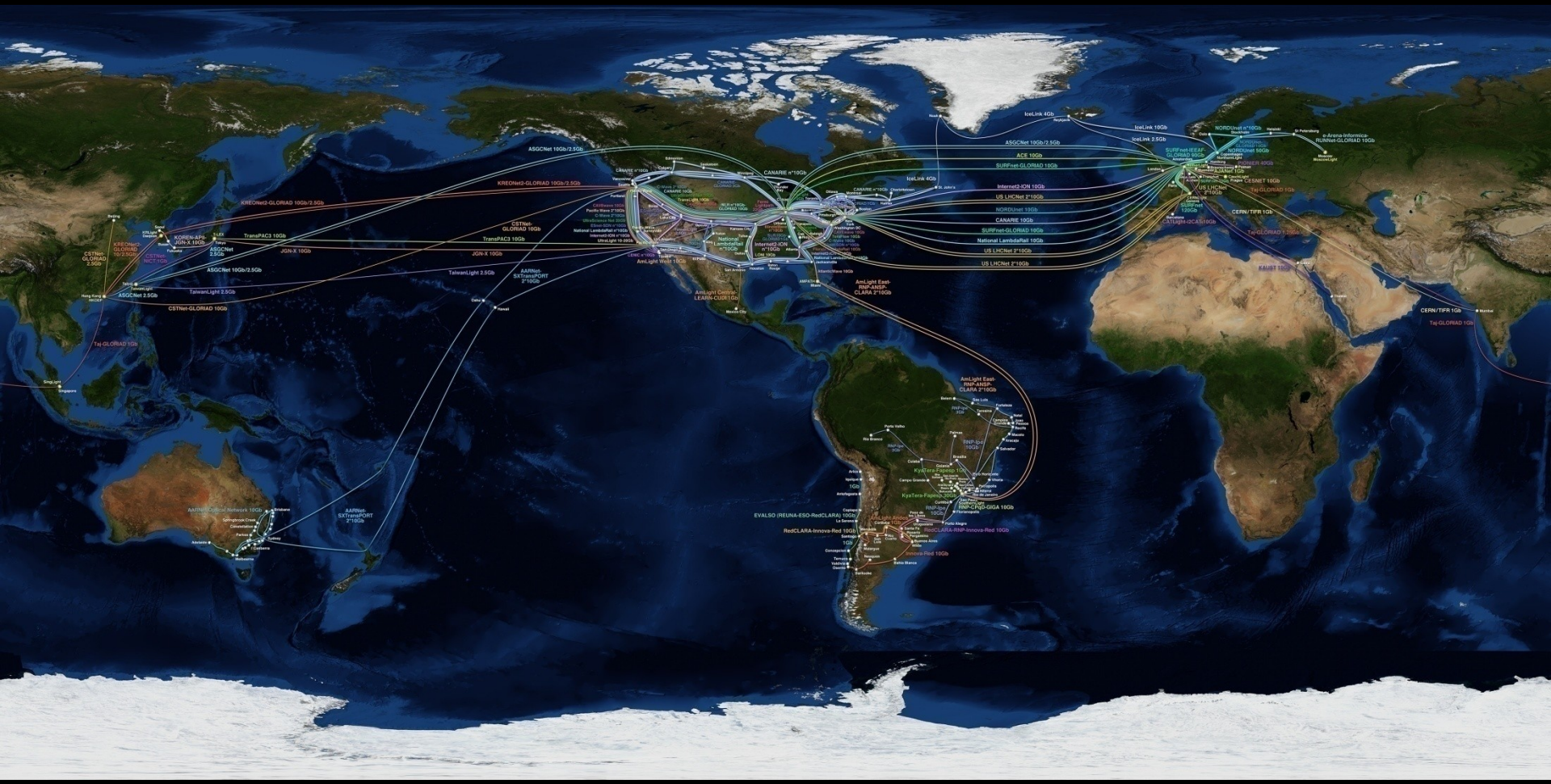
This book, edited by four of the leaders of the National Science Foundation's Global Environment and Network Innovations (GENI) project, gives the reader a tour of the history, architecture, future, and applications of GENI. Built over the past decade by hundreds of leading computer scientists and engineers, GENI is a nationwide network used daily by thousands of computer scientists to explore the next Cloud and Internet and the applications and services they enable, which will transform our communities and our lives. Since by design it runs on existing computing and networking equipment and over the standard commodity Internet, it is poised for explosive growth and transformational impact over the next five years.

iGENI: The International GENI

- The iGENI Initiative Will Design, Develop, Implement, and Operate a Major New National and International Distributed Infrastructure.
- iGENI Will Place the "G" in GENI Making GENI Truly Global.
- iGENI Will Be a Unique Distributed Infrastructure Supporting Research and Development for Next-Generation Network Communication Services and Technologies.
- This Infrastructure Will Be Integrated With Current and Planned GENI Resources, and Operated for Use by GENI Researchers Conducting Experiments that Involve Multiple Aggregates At Multiple Sites.
- iGENI Infrastructure Will Connect Its Resources With Current GENI National Backbone Transport Resources, With Current and Planned GENI Regional Transport Resources, and With International Research Networks and Projects,
-

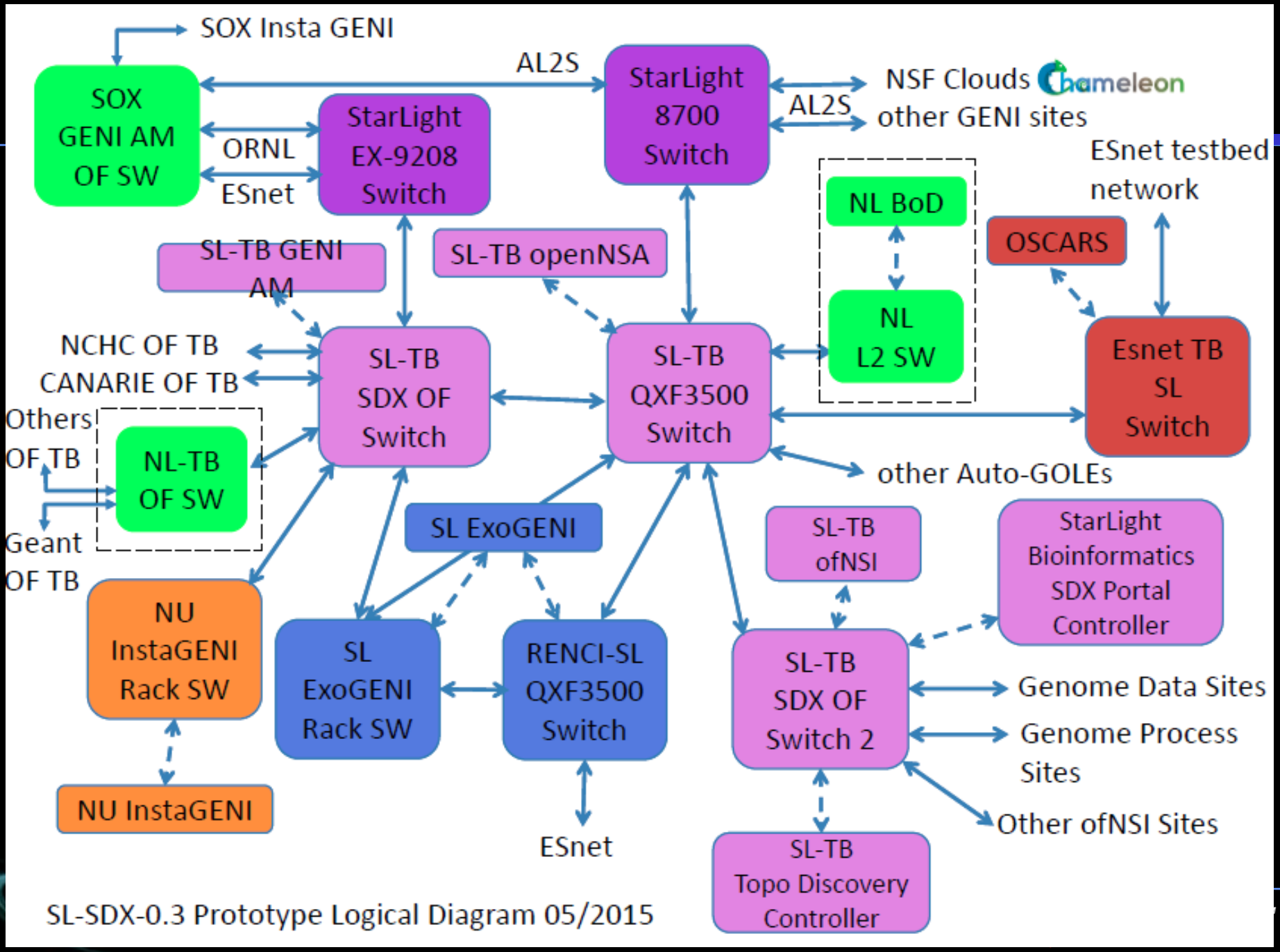

KDDI
NSI
AIST
OpenFlow
NSI
KUAS
NetFPGA
OpenFlow
NCKU
NetFPGA
OpenFlow
TOMATO
OpenFlow
InstaGENI
University of Victoria
ciena
ExoGENI
OpenFlow
NSI
cenic
InstaGENI
OpenFlow
JGN-X
NSI
東京大学
THE UNIVERSITY OF TOKYO
VNode
OpenFlow
OpenFlow
ETRI
OpenFlow
KiSTi
NCHC
NSI
NetFPGA
OpenFlow
ofNSI
KRLight
NSI
canarie
OpenFlow
PACIFIC NORTHWEST GIGAPOP
aarnet
NICTA
ExoGENI
OpenFlow
APPLIED RESEARCH CENTER FOR COMPUTER NETWORKS
OpenFlow
GLORIAD
NSI
PACIFIC WAVE
GRPNet
ASGC
OSDC
OpenFlow
GT
InstaGENI
OpenFlow
CLEMSON UNIVERSITY
TOMATO
InstaGENI
OpenFlow
SingAREN
OpenFlow
NetFPGA
SingLight
MREN
STARLIGHT
The Optical STAR TAP
GENI AM
ExoGENI
OpenFlow
NSI
ofNSI
OSDC
ASGCNet
NSI
MAX
MID-ATLANTIC CROSSROADS
GENI AM
InstaGENI
OpenFlow
SoX
Southern Crossroads
GENI AM
InstaGENI
OpenFlow
renci
ExoGENI
OpenFlow
CNC
OpenFlow
NORTHWESTERN UNIVERSITY
NetFPGA
OpenFlow
TOMATO
InstaGENI
THE UNIVERSITY OF CHICAGO
OSDC
OpenFlow
INTERNET2
GENI AM
OpenFlow
BEN
Boston Educational Network
ExoGENI
Duke UNIVERSITY
ExoGENI
OpenFlow
OpenFlow
AMPATH
NSI
OpenFlow
CPqD
OpenFlow
RNP
NSI
NL Light
NSI
OpenFlow
NetFPGA
MAGPI
UNIVERSITY OF AMSTERDAM
NSI
ExoGENI
Princeton University
OpenFlow
InstaGENI
uen
UTAH EDUCATION NETWORK
GENI AM
OpenFlow
THE UNIVERSITY OF UTAH
TOMATO
VNode
OpenFlow
InstaGENI
USP
InstaGENI
TOMATO
TECHNISCHE UNIVERSITÄT KAISERSLAUTERN
DFN
Deutsches Forschungsnetz
GÉANT
NSI
ja.net
University of BRISTOL
OpenFlow
University of Essex
OpenFlow
grnet
NSI
POZNAN
OpenFlow
NSI
Belnet
dedicated connectivity
iMinds
InstaGENI

The Global Lambda Integrated Facility: a Global Programmable Resource

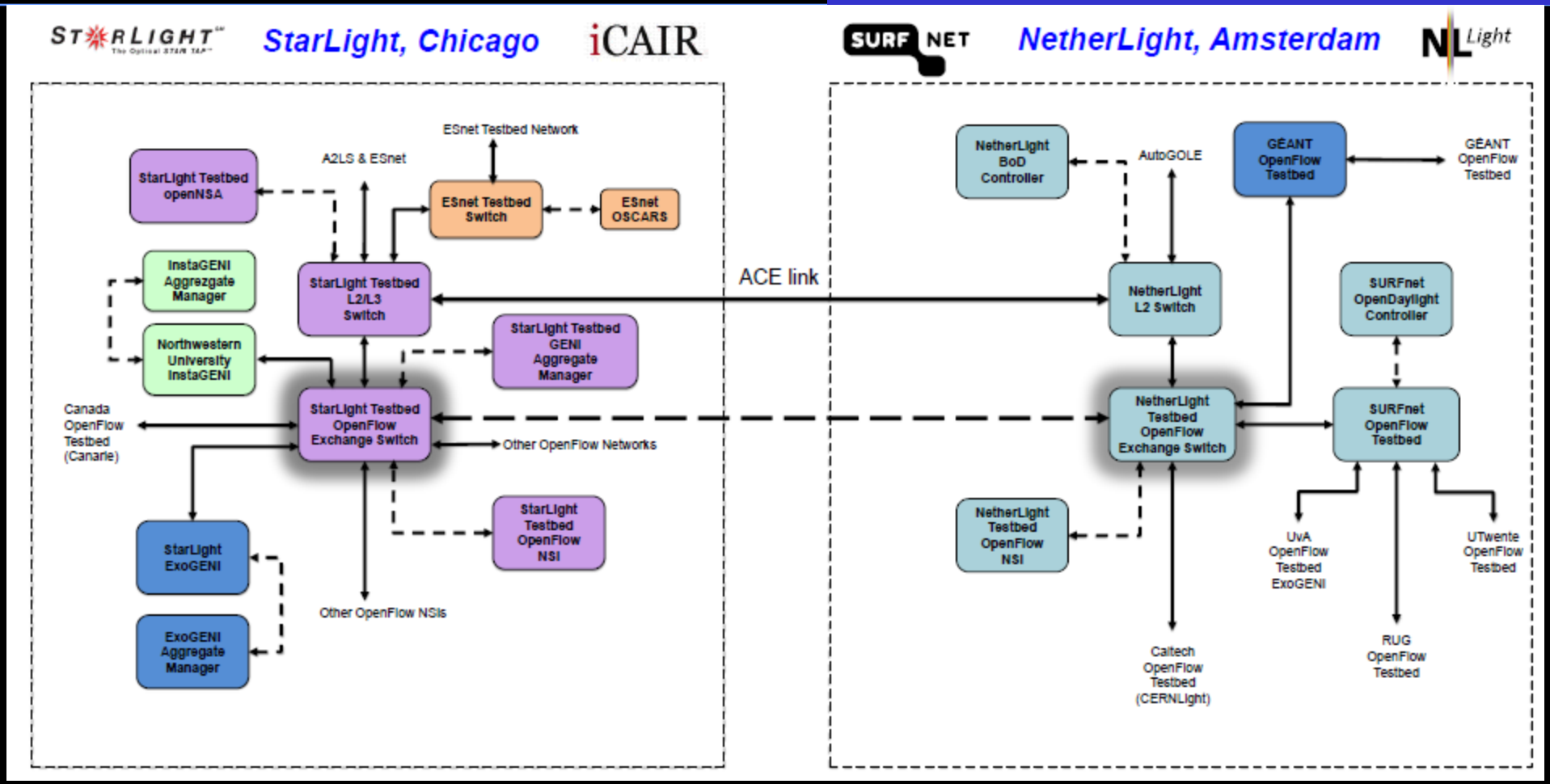


iCAIR

STARLIGHT SDX


SOX Insta GENI
SOX GENI AM OF SW
AL2S
StarLight 8700 Switch
NSF Clouds
Chameleon
AL2S
other GENI sites
ORNL
ESnet
StarLight EX-9208 Switch
ESnet testbed network
NL BoD
OSCARS
SL-TB GENI AM
SL-TB openNSA
NL L2 SW
NCHC OF TB
CANARIE OF TB
SL-TB SDX OF Switch
SL-TB QXF3500 Switch
Esnet TB SL Switch
Others OF TB
NL-TB OF SW
other Auto-GOLEs
Geant OF TB
SL ExoGENI
SL-TB ofNSI
StarLight Bioinformatics SDX Portal Controller
NU InstaGENI Rack SW
SL ExoGENI Rack SW
RENCI-SL QXF3500 Switch
SL-TB SDX OF Switch 2
Genome Data Sites
Genome Process Sites
Other ofNSI Sites
NU InstaGENI
ESnet
SL-TB Topo Discovery Controller
SL-SDX-0.3 Prototype Logical Diagram 05/2015

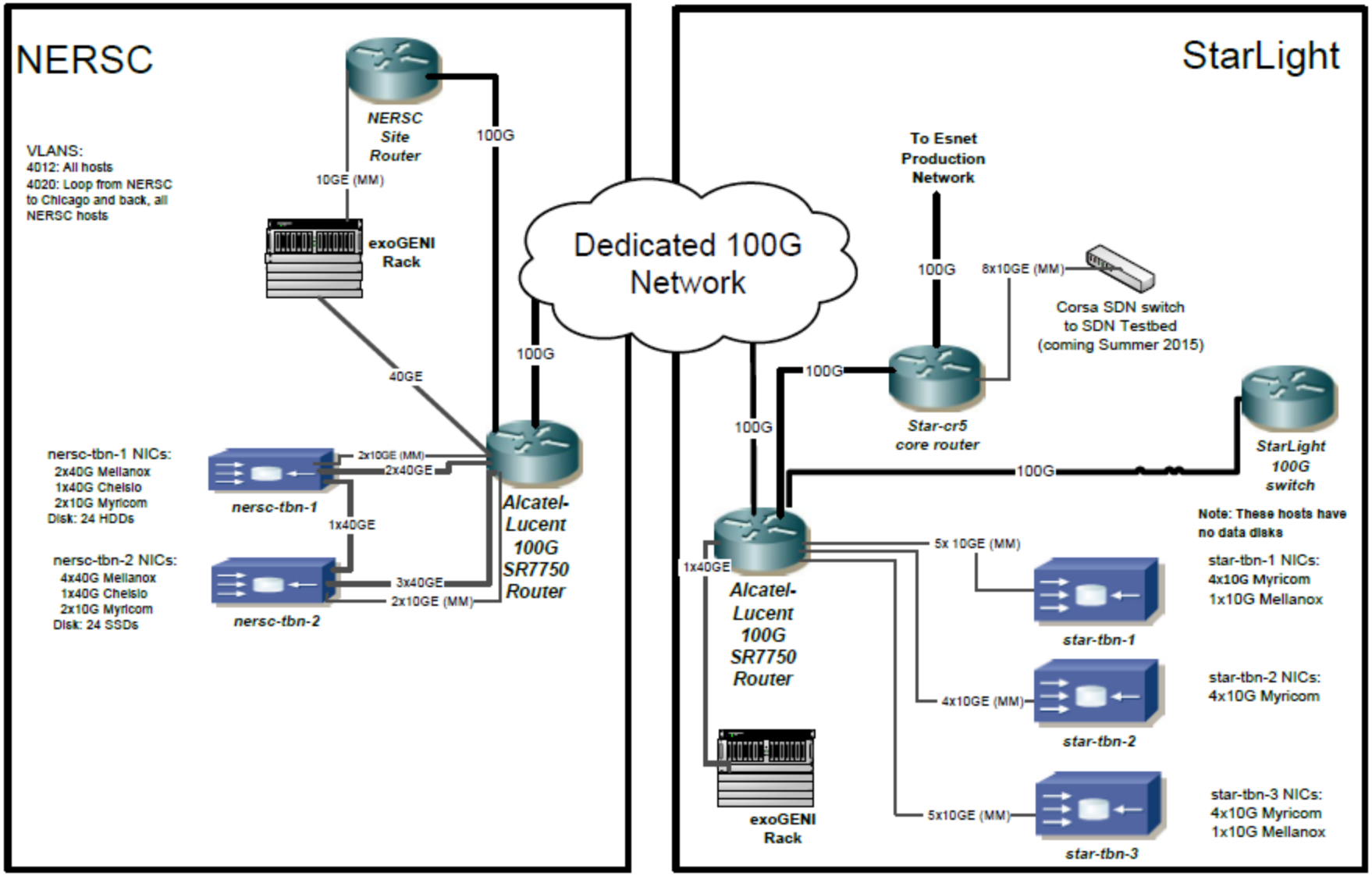
SDX StarLight⇔NetherLight



Ronald van der Pol, Joe Mambretti, Jim Chen, John Shillington

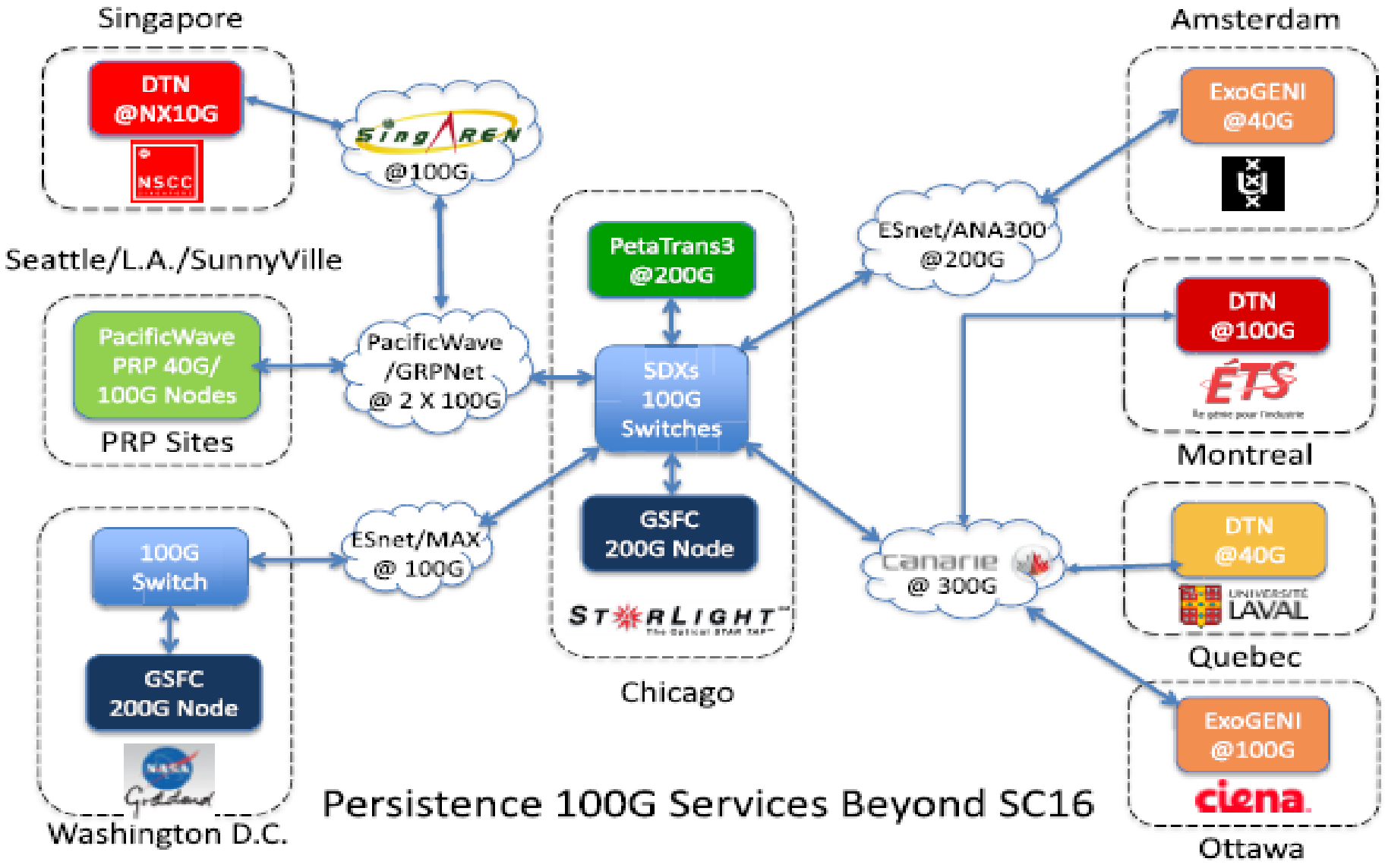
STARLIGHT SM

100G Component of Esnet SDN Testbed



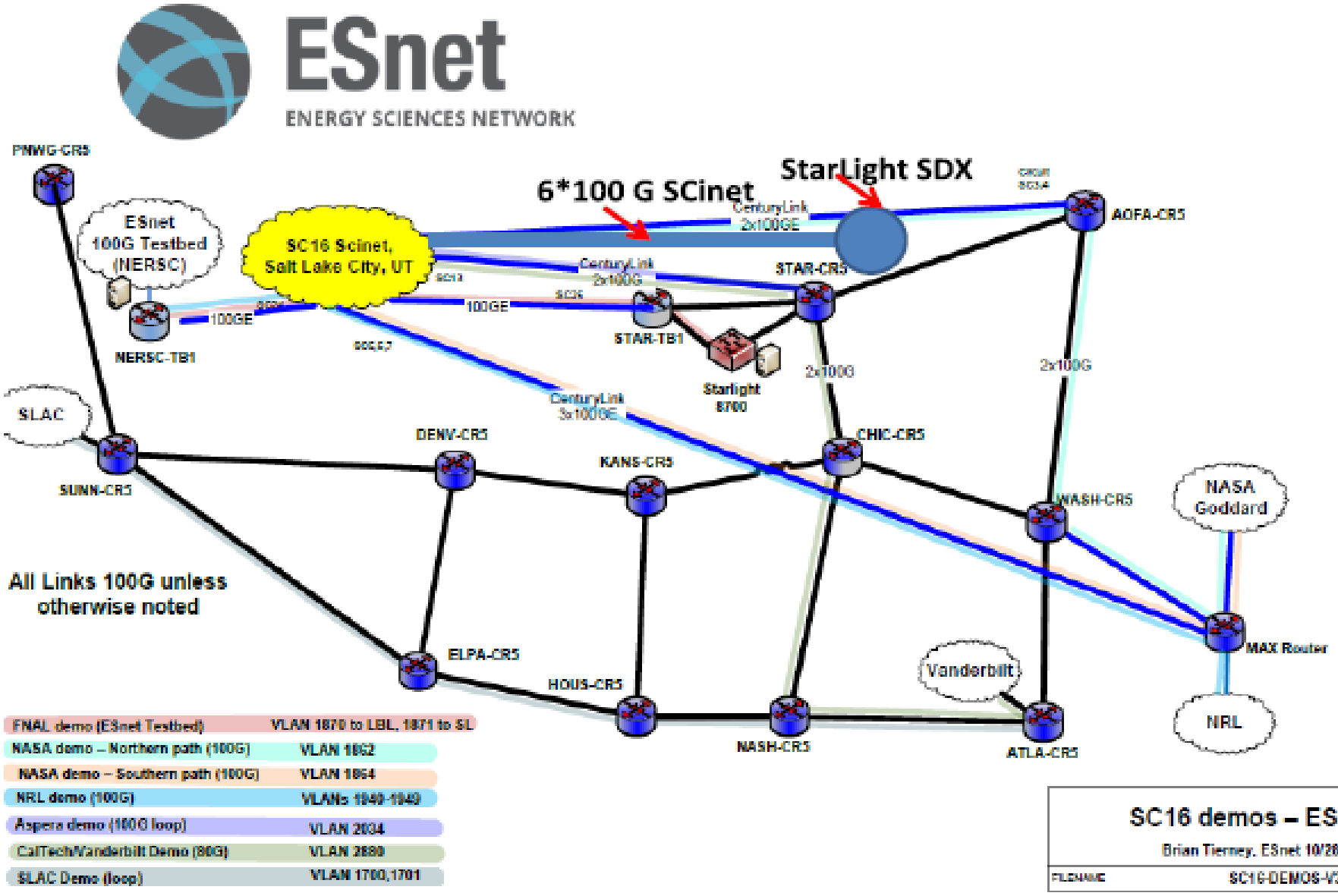
STARLIGHT SM

PetaTrans: Petascale Sciences Data Transfer

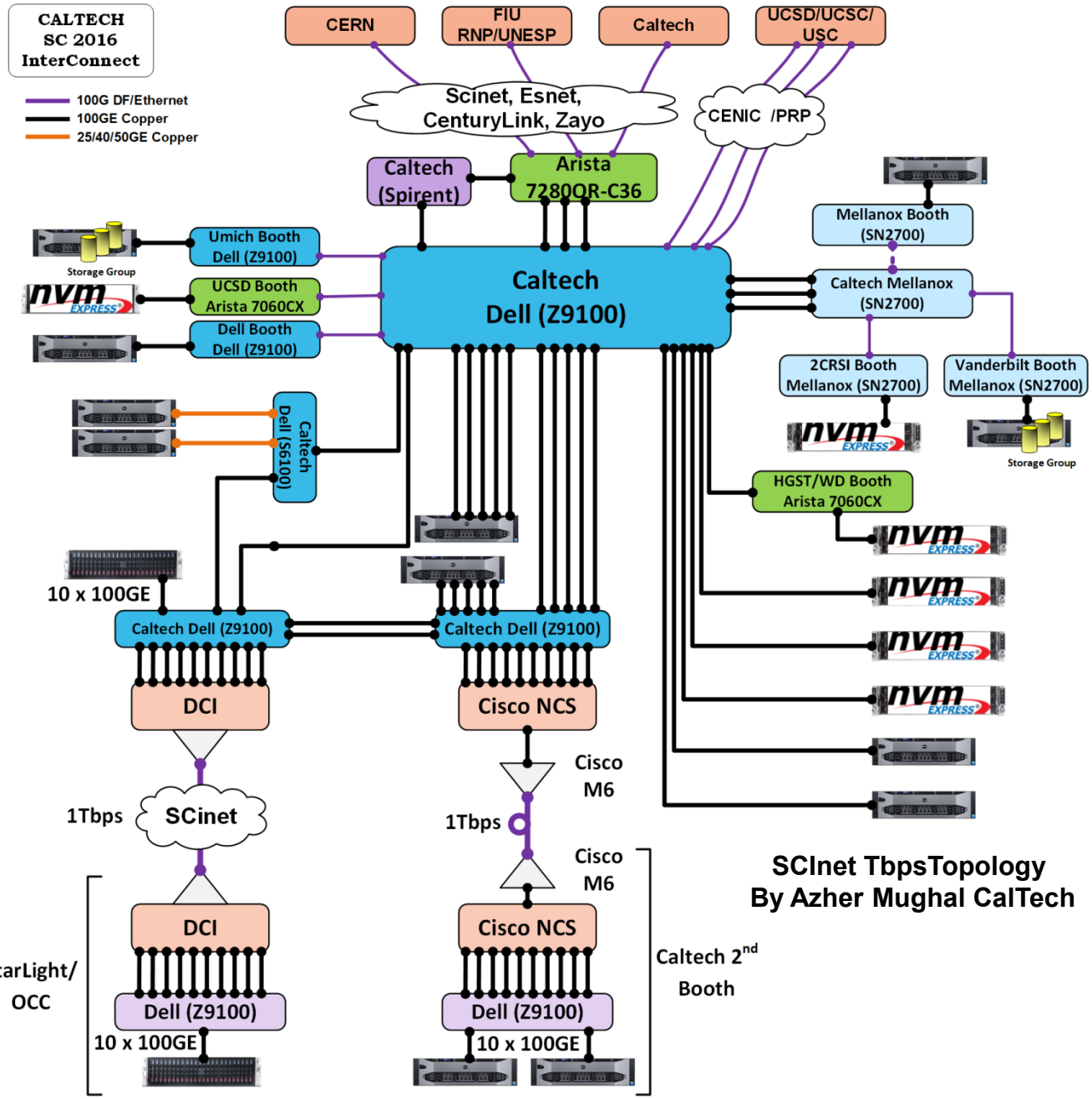


Persistence 100G Services Beyond SC16

STARLIGHT℠



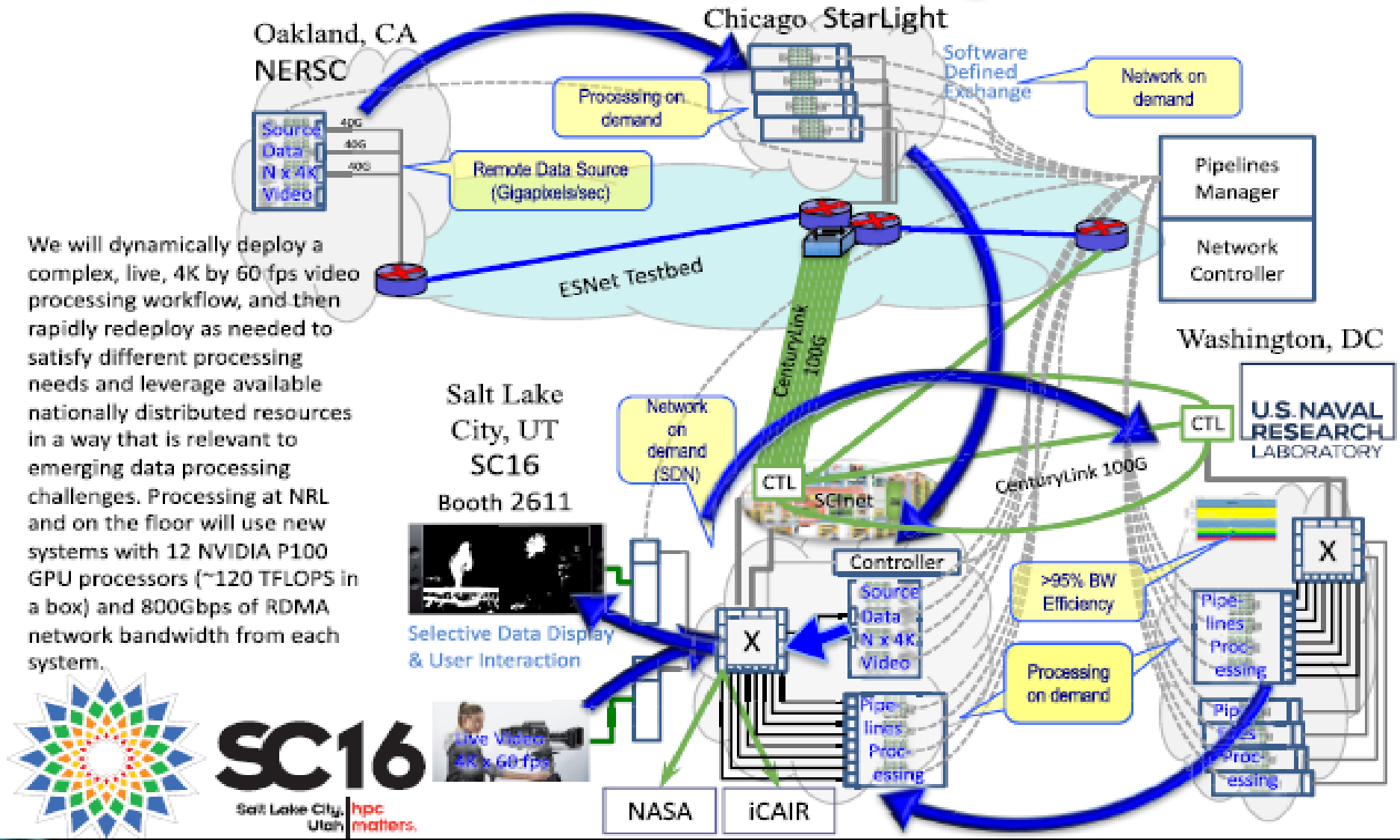
STARLIGHT SM


CALTECH
SC 2016
InterConnect
100G DF/Ethernet
100GE Copper
25/40/50GE Copper
CERN
FIU
RNP/UNESP
Caltech
UCSD/UCSC/
USC
Scinet, Esnet,
CenturyLink, Zayo
CENIC /PRP
Caltech
(Spirent)
Arista
7280QR-C36
Mellanox Booth
(SN2700)
Storage Group
Umich Booth
Dell (Z9100)
UCSD Booth
Arista 7060CX
Dell Booth
Dell (Z9100)
Caltech
Dell (Z9100)
Caltech Mellanox
(SN2700)
2CRSI Booth
Mellanox (SN2700)
Vanderbilt Booth
Mellanox (SN2700)
Storage Group
Caltech
Dell (S6100)
HGST/WD Booth
Arista 7060CX
10 x 100GE
Caltech Dell (Z9100)
Caltech Dell (Z9100)
DCI
Cisco NCS
Cisco
M6
1Tbps
SCinet
1Tbps
Cisco
M6
DCI
Cisco NCS
StarLight/
OCC
Dell (Z9100)
Dell (Z9100)
10 x 100GE
10 x 100GE
Caltech 2nd
Booth
SCInet TbpsTopology
By Azher Mughal CalTech

IGHT SM

Dynamic Remote I/O

for Distributed Compute and Storage



We will dynamically deploy a complex, live, 4K by 60 fps video processing workflow, and then rapidly redeploy as needed to satisfy different processing needs and leverage available nationally distributed resources in a way that is relevant to emerging data processing challenges. Processing at NRL and on the floor will use new systems with 12 NVIDIA P100 GPU processors (~120 TFLOPS in a box) and 800Gbps of RDMA network bandwidth from each system.

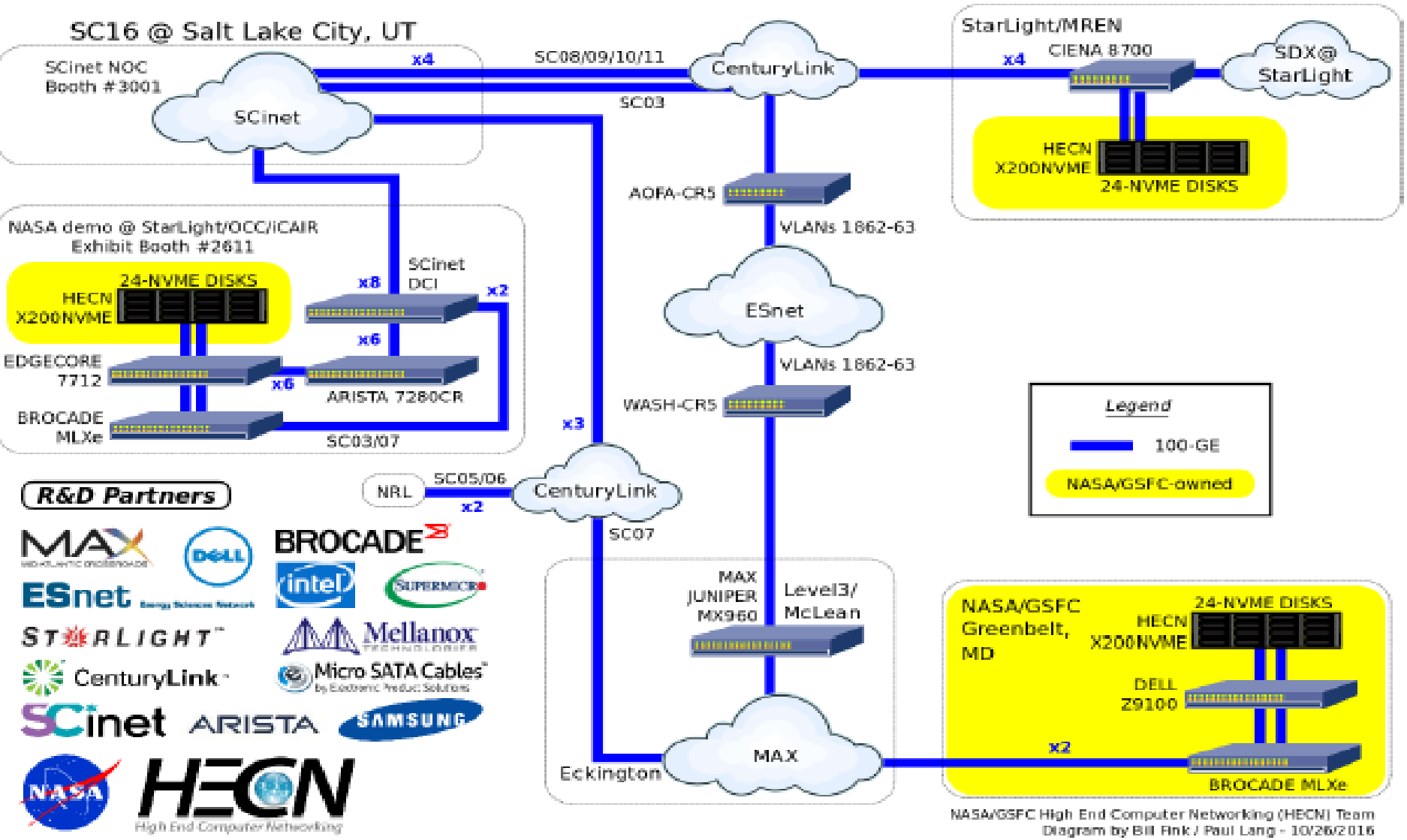
SC16 Salt Lake City, Utah hpc matters.

STARLIGHT SM

SC16

Demonstrations of 200 Gbps Disk-to-Disk WAN File Transfers using Parallelism across NVMe Drives

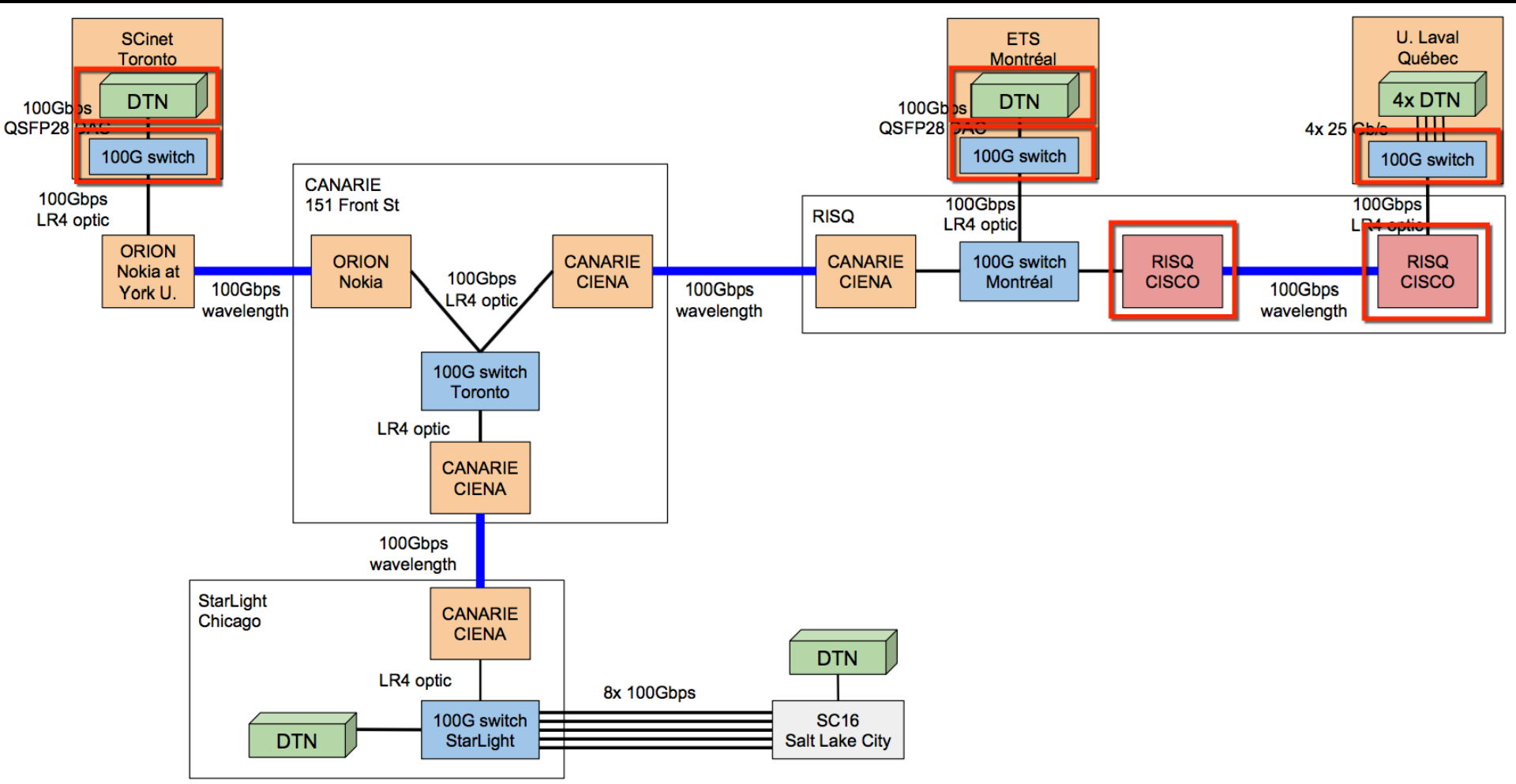
An SC16 Collaborative Initiative Among NASA and Several Partners



NASA/GSFC High End Computer Networking (HECN) Team
Diagram by Bill Fink / Paul Lang - 10/26/2016

SC16 SDN-WAN Demonstration End-Points
Caltech, UM, Vanderbilt, UCSD, Dell, 2CRSI, KISTI, StarLight, PRP, FIU, RNP, UNESP, CERN

DTN Flows@100 Gbps=>Compute Canada⇔CANARIE⇔StarLight<+>SC16



mdtmFTP @ 100Gbps Networks

mdtmFTP: a high-performance data transfer tool

- Pipelined I/O-centric design to streamline data transfer
- Multicore-aware data transfer middleware (MDTM) optimizes use of underlying multicore system
- Extremely efficient in transferring of Lots Of Small Files (LOSF)

Fermilab

ESnet ENERGY SCIENCES NETWORK

STARLIGHT SM The Optical STAR TAP SM

iCAIR

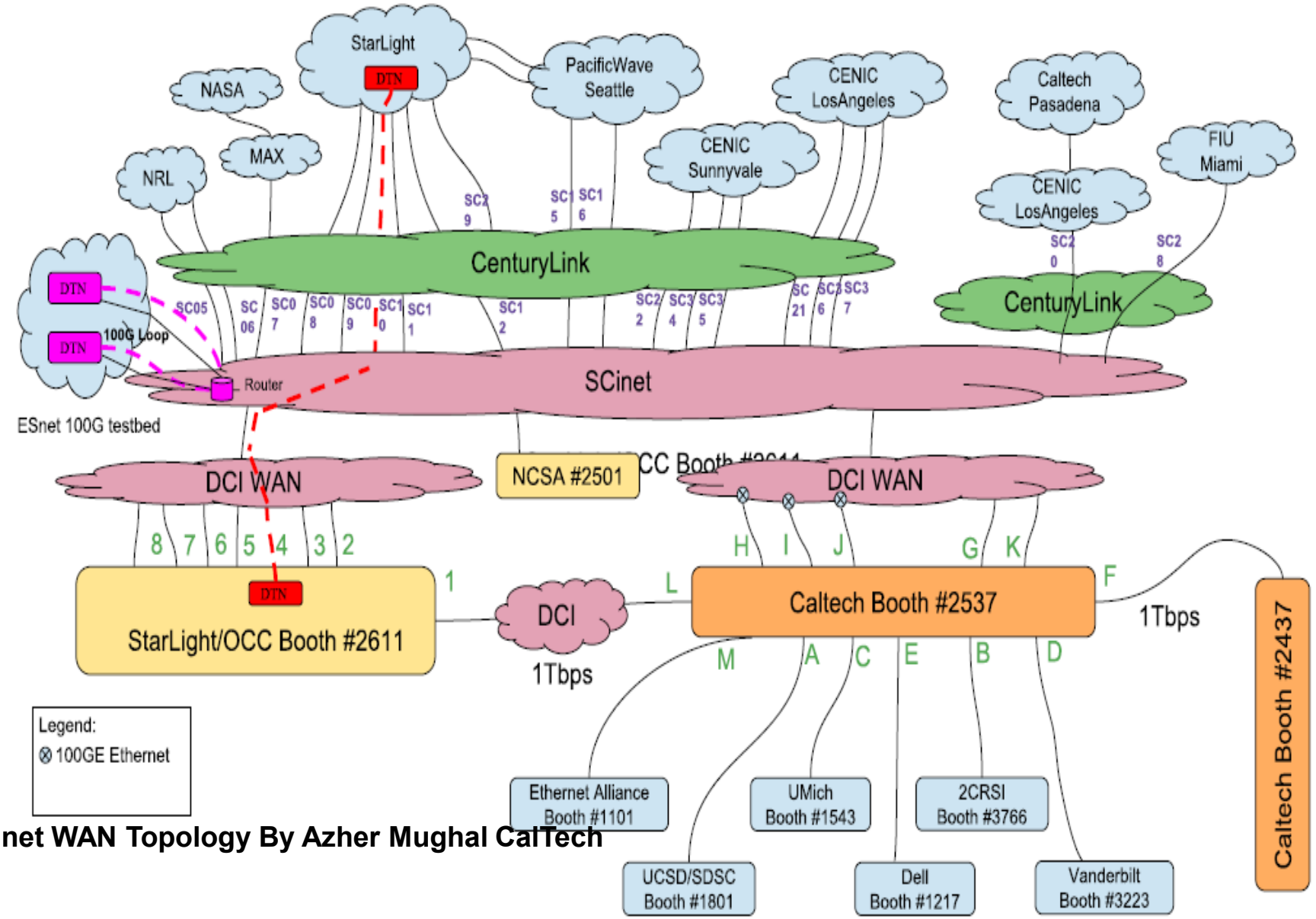
NERSC

NARLabs 國家實驗研究院 National Applied Research Laboratories

A DOE/SC/ASCR-sponsored research project

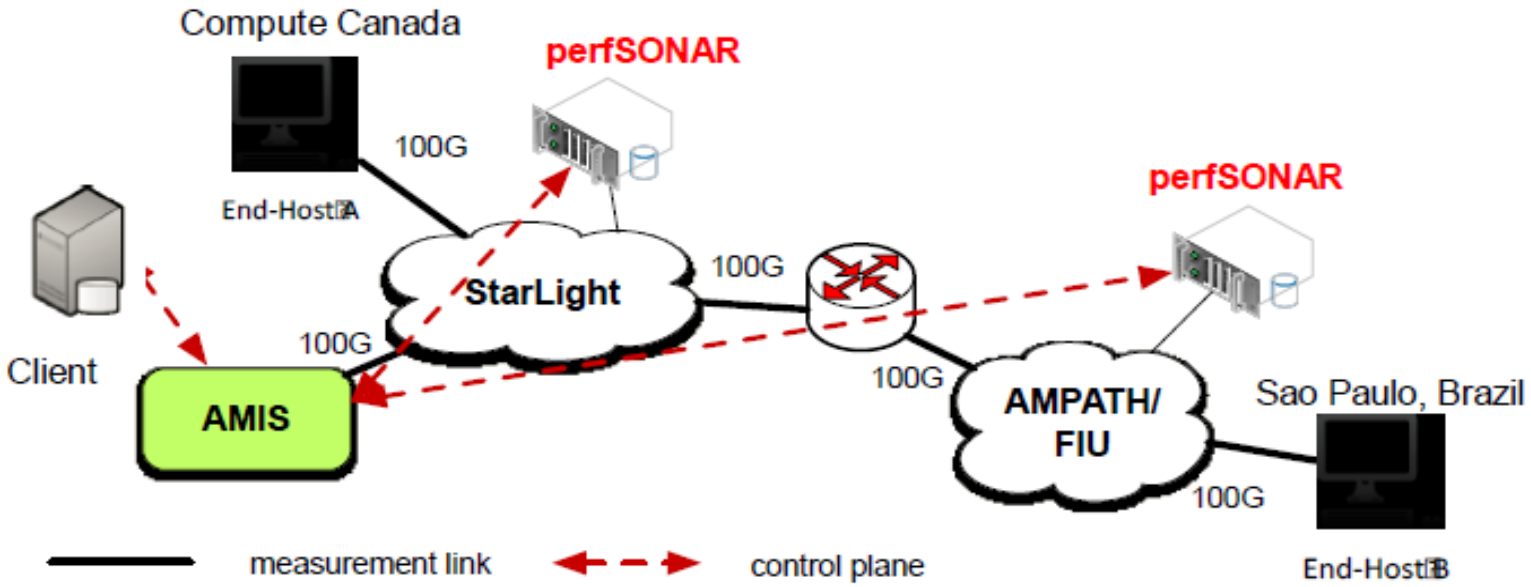
Software is available at: http://mdtm.fnal.gov

mdtmFTP @ 100Gbps Networks



SCInet WAN Topology By Azher Mughal CalTech

Programmable Network Measurement of Data Intensive Flows on 100Gbps Networks



Demo1: Programmable Measurement with RESTful APIs

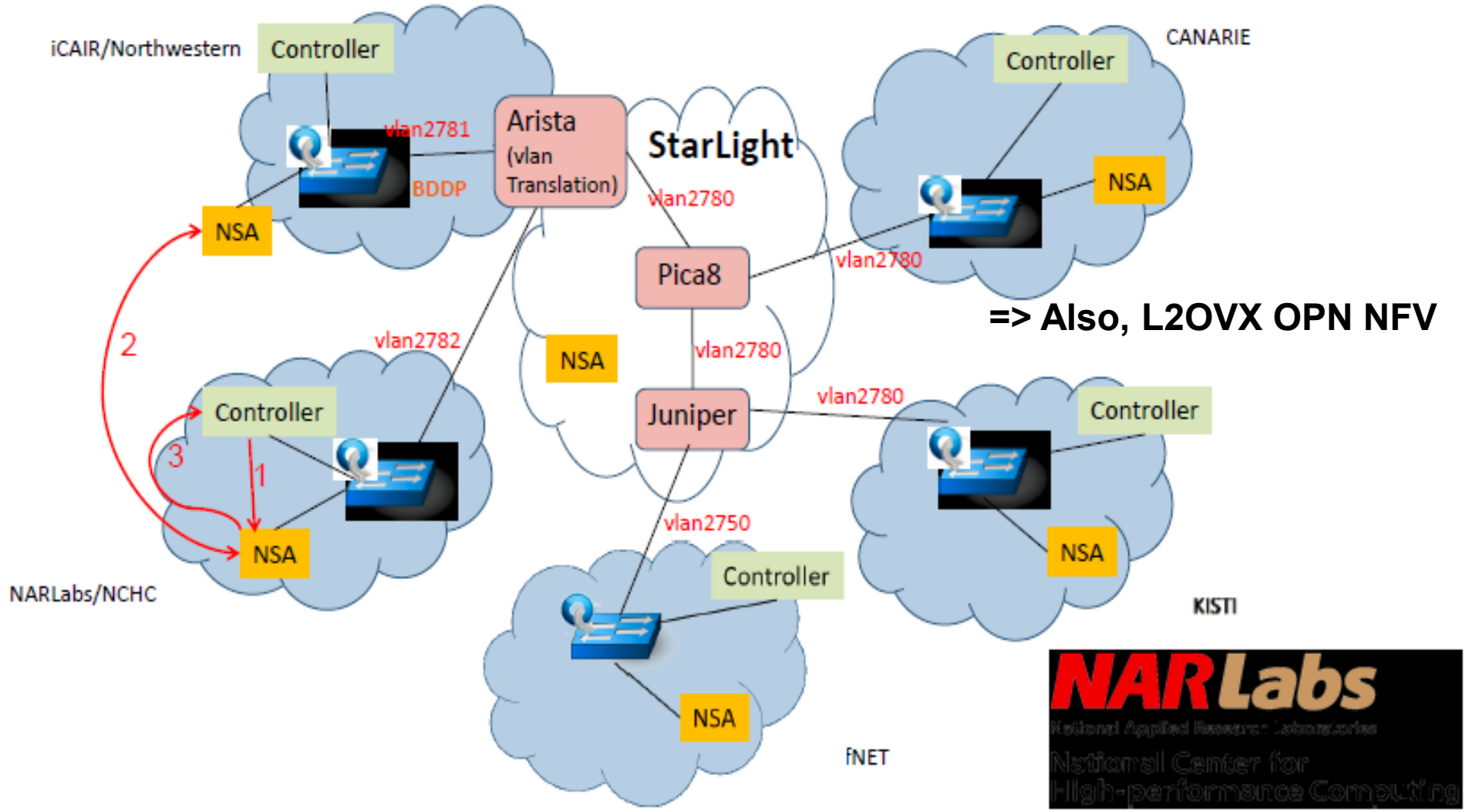
Demo2: Passive & Active Measurement (TCP window size)

Demo3: Passive & Active Measurement (TCP packet loss)

RNC AMIS Team: Yan Luo, PI, University of Massachusetts Lowell; Gabriel Ghinita, Co-PI, University of Massachusetts Boston; Cody Bumgardner, Co-PI, University of Kentucky; Michael McGarry, Co-PI, University of Texas El Paso. Contact: **Yan_Luo@uml.edu**

Collaborators: Jeo Mambretti, Jim Chen and Fei Yeh, StarLight/iCAIR/Northwestern University; Jeronimo Bezerra, AMPATH/Florida International University

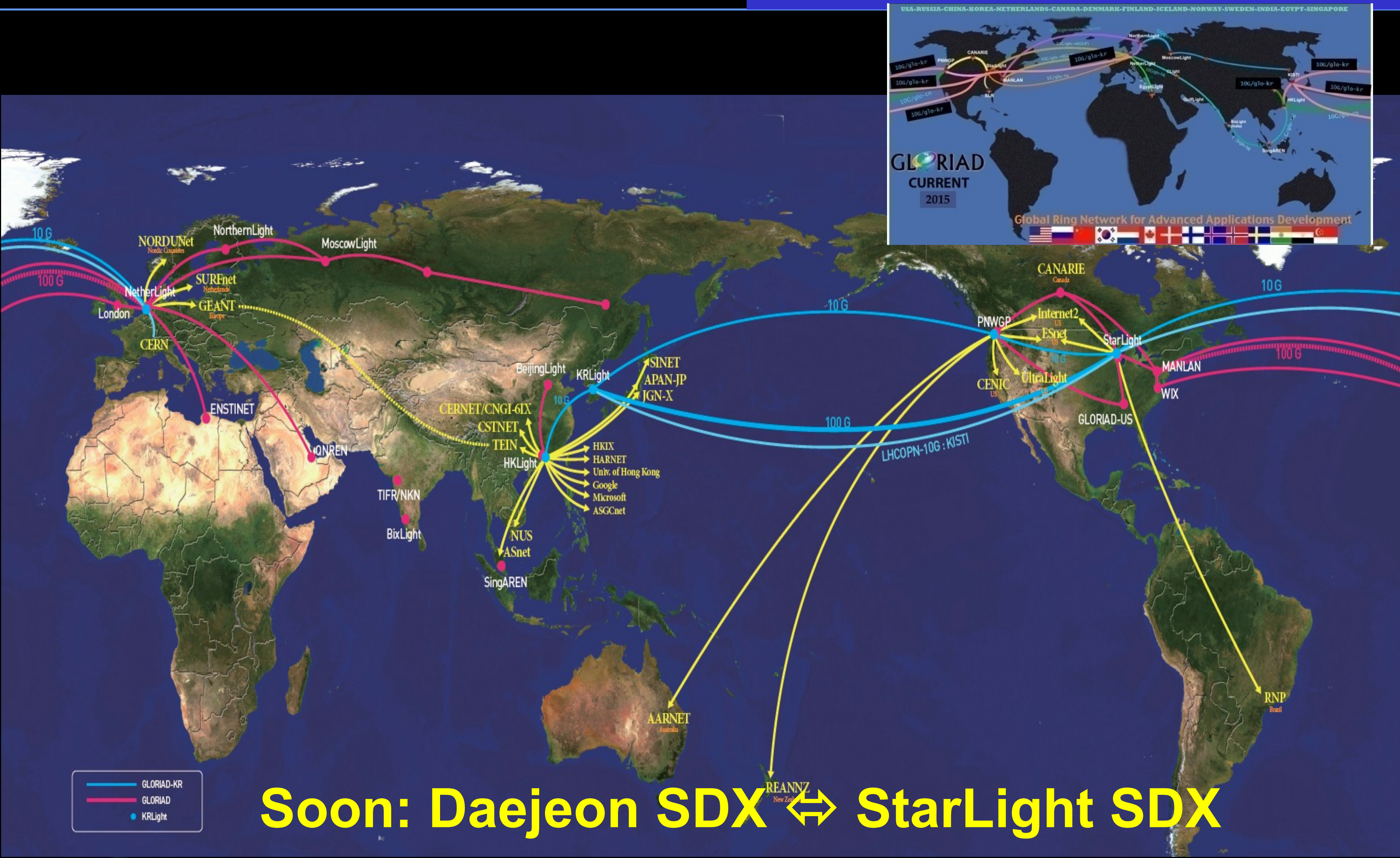
NSI-OpenFlow Hybird Topology Exchange



Coming Soon =>Taiwan SDX

STARLIGHT℠

KREONet2 and GLORIAD-KR
KISTI Daejeon ⇔ 100 G ⇔ StarLight



Soon: Daejeon SDX ⇔ StarLight SDX

Nationwide 100 Gbps and Minimized Latency

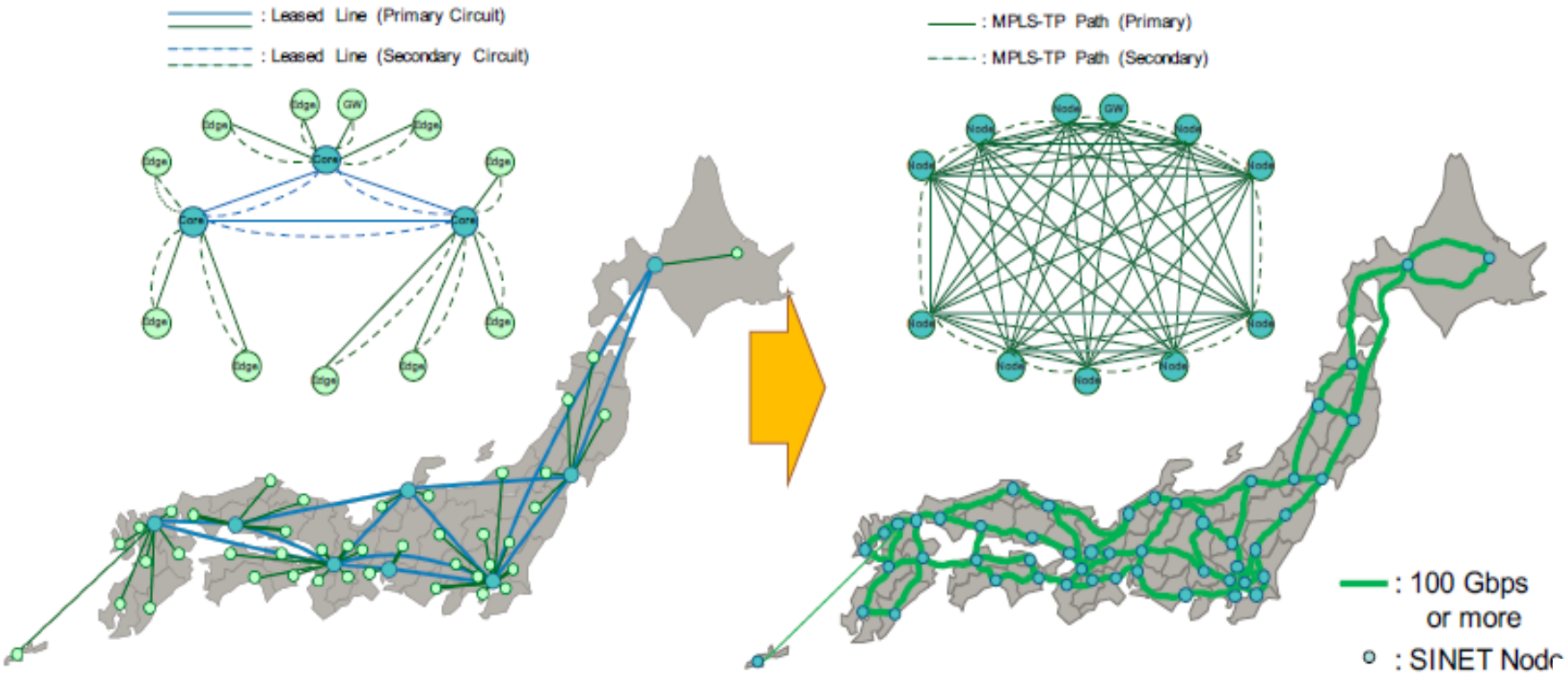
◆ SINET5 will be a nationwide 100-Gbps backbone network using 100-Gigabit Ethernet technology and connect each pair of nodes with a minimized latency.

SINET4

- Star-like topology
- Resource-consuming secondary circuits

SINET5

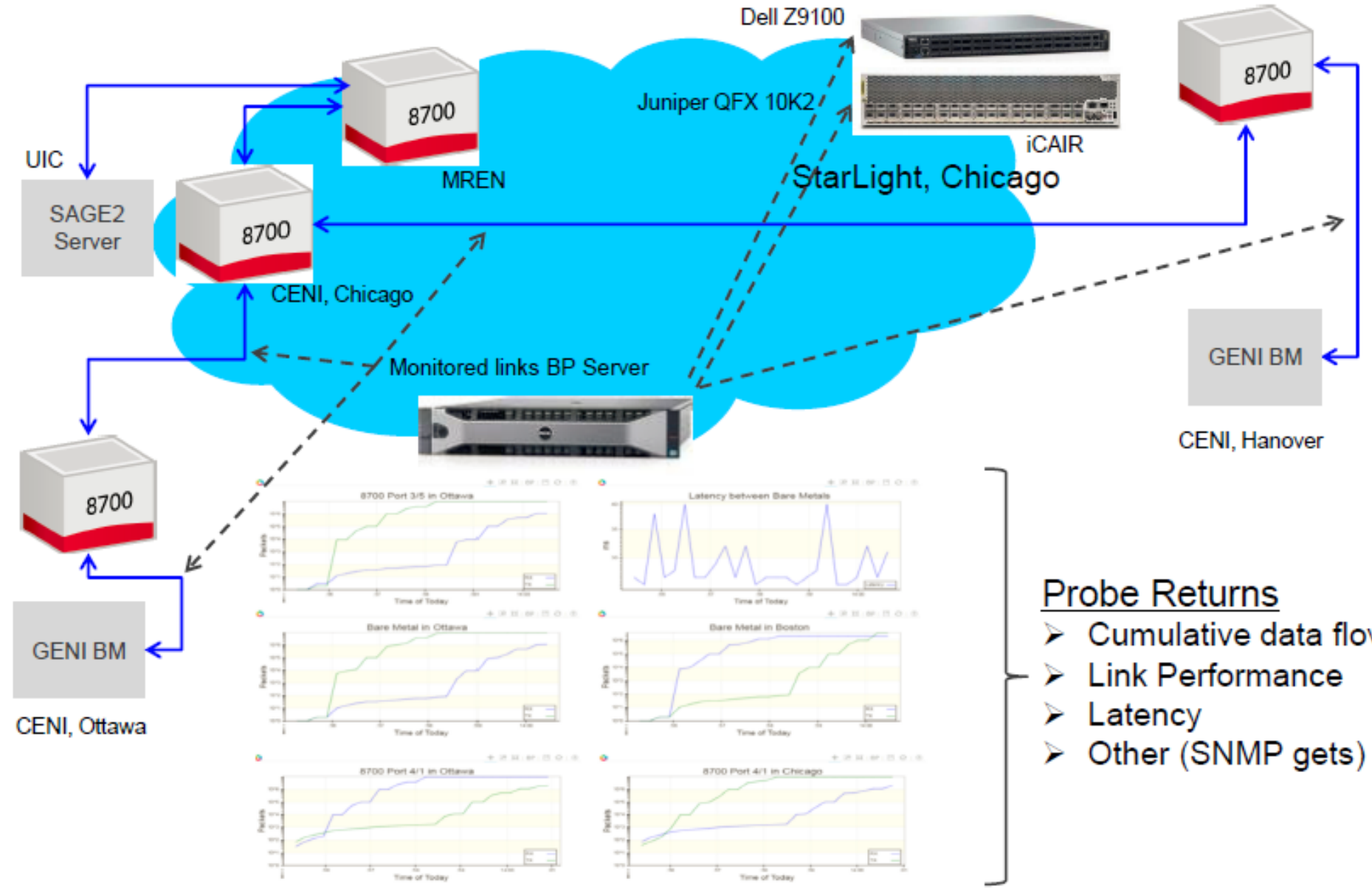
- Fully-meshed topology with redundancy
- Non-resource-consuming secondary paths



A*STAR Singapore

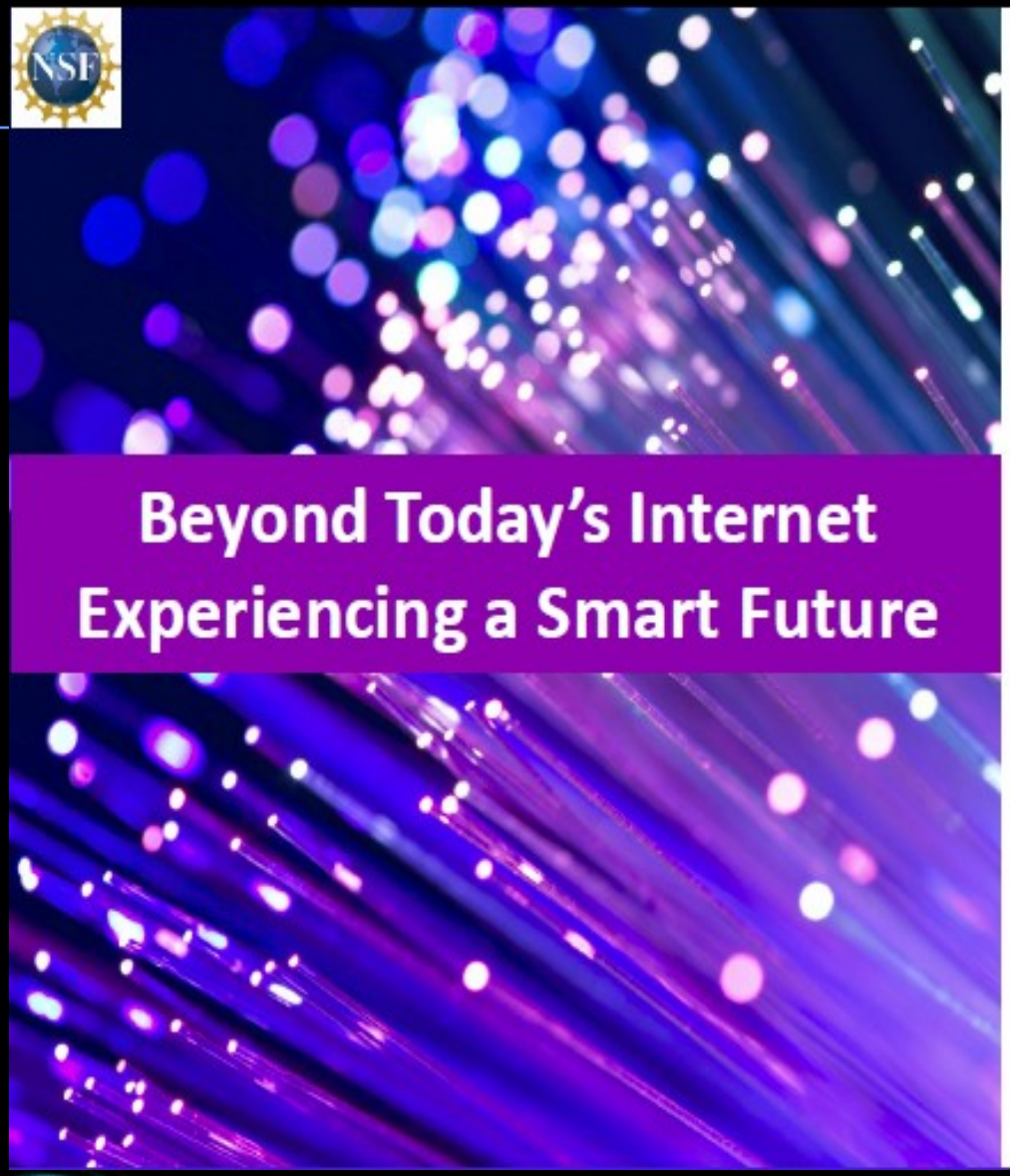
- Singapore Supercomputing Center DTN⇔SingAREN⇔PacWavw⇔GRPnet⇔ StarLight DTN⇔SC16
- 50-60 Gbps

Network Diagram for Analytics Demonstration



Probe Returns

- Cumulative data flows
- Link Performance
- Latency
- Other (SNMP gets)

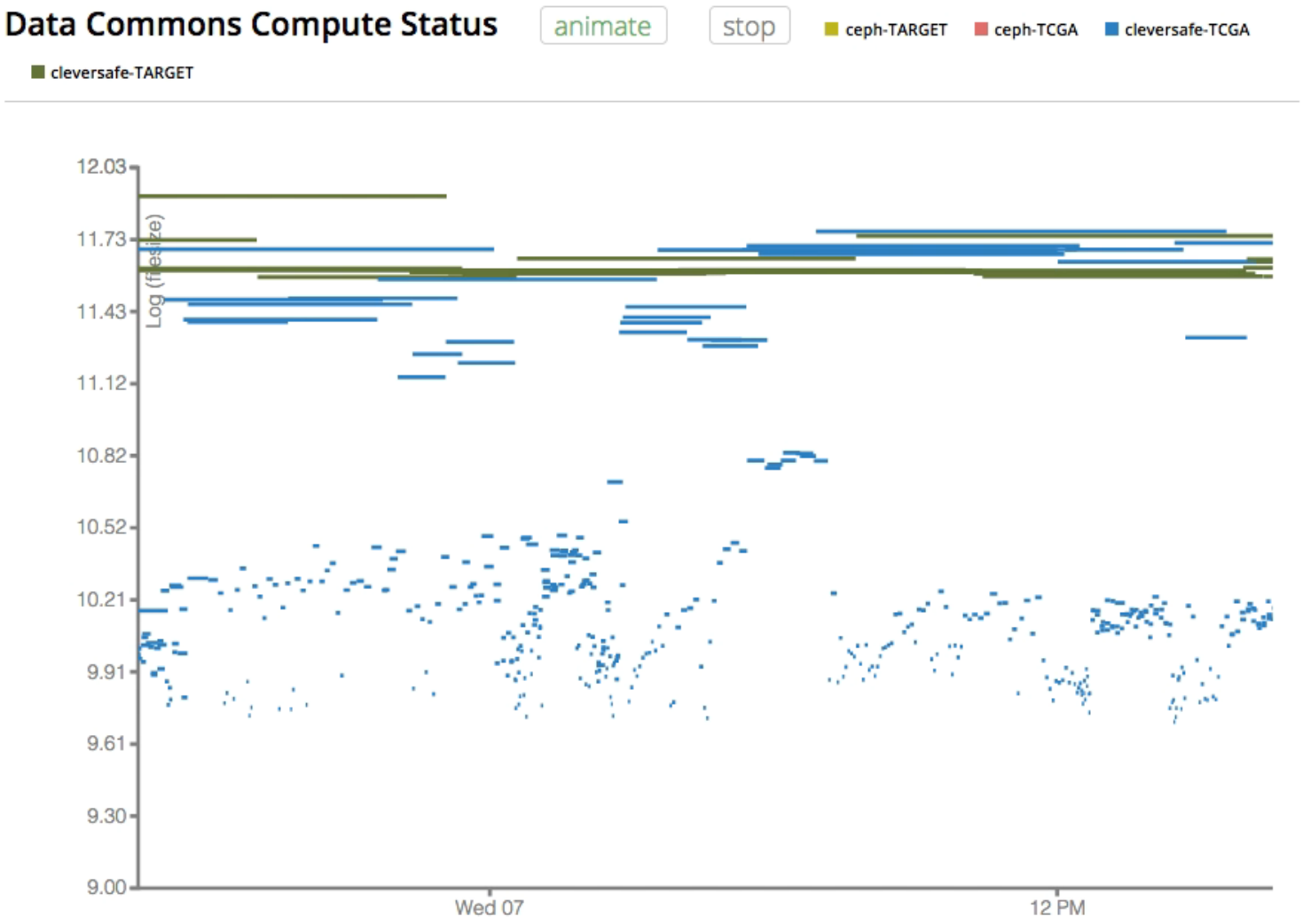


Prototype SDX Bioinformatics Exchange: Demonstrating an Essential Use-Case for Personalized Medicine

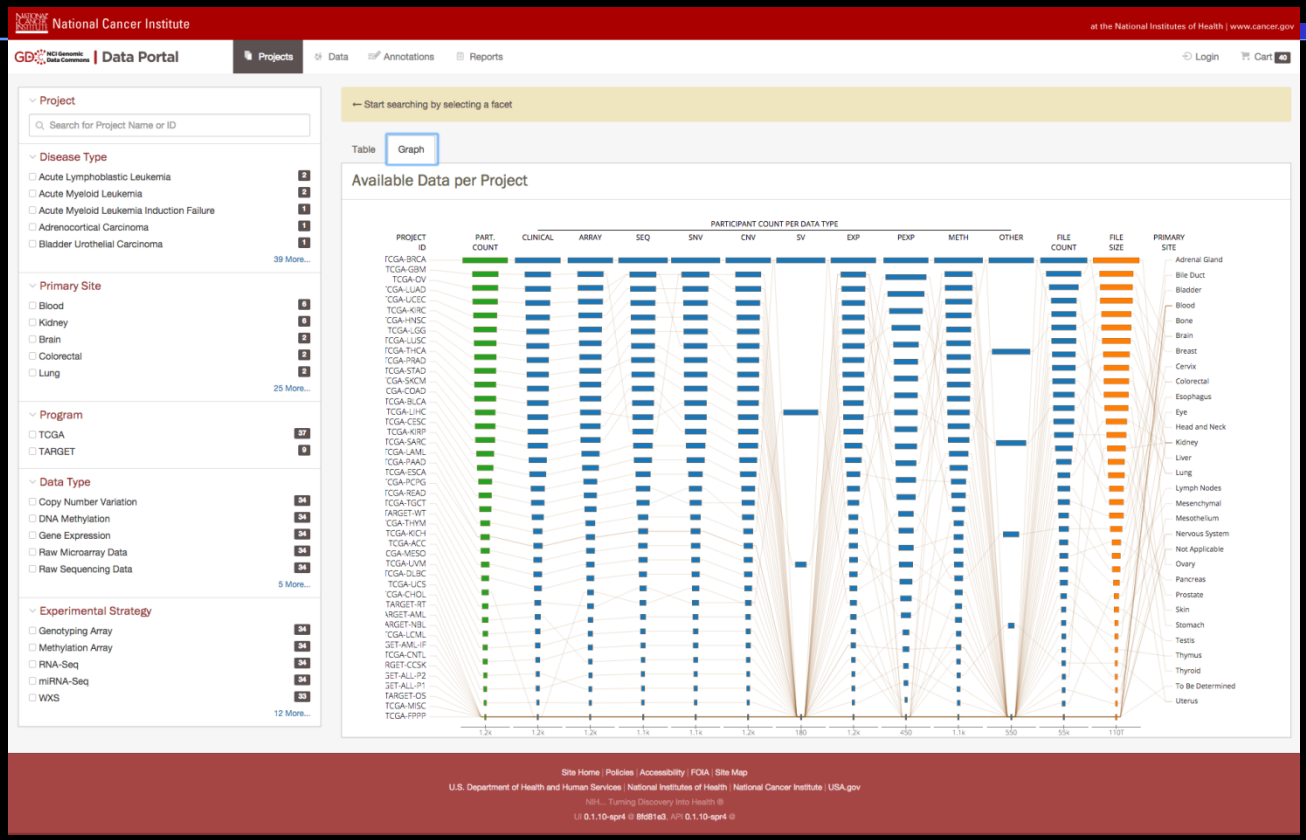
Robert Grossman, Piers Nash, Allison Heath, Renuka Arya
University of Chicago

Joe Mambretti, Jim Chen
Northwestern University

Genomic Data Commons Data Transfer



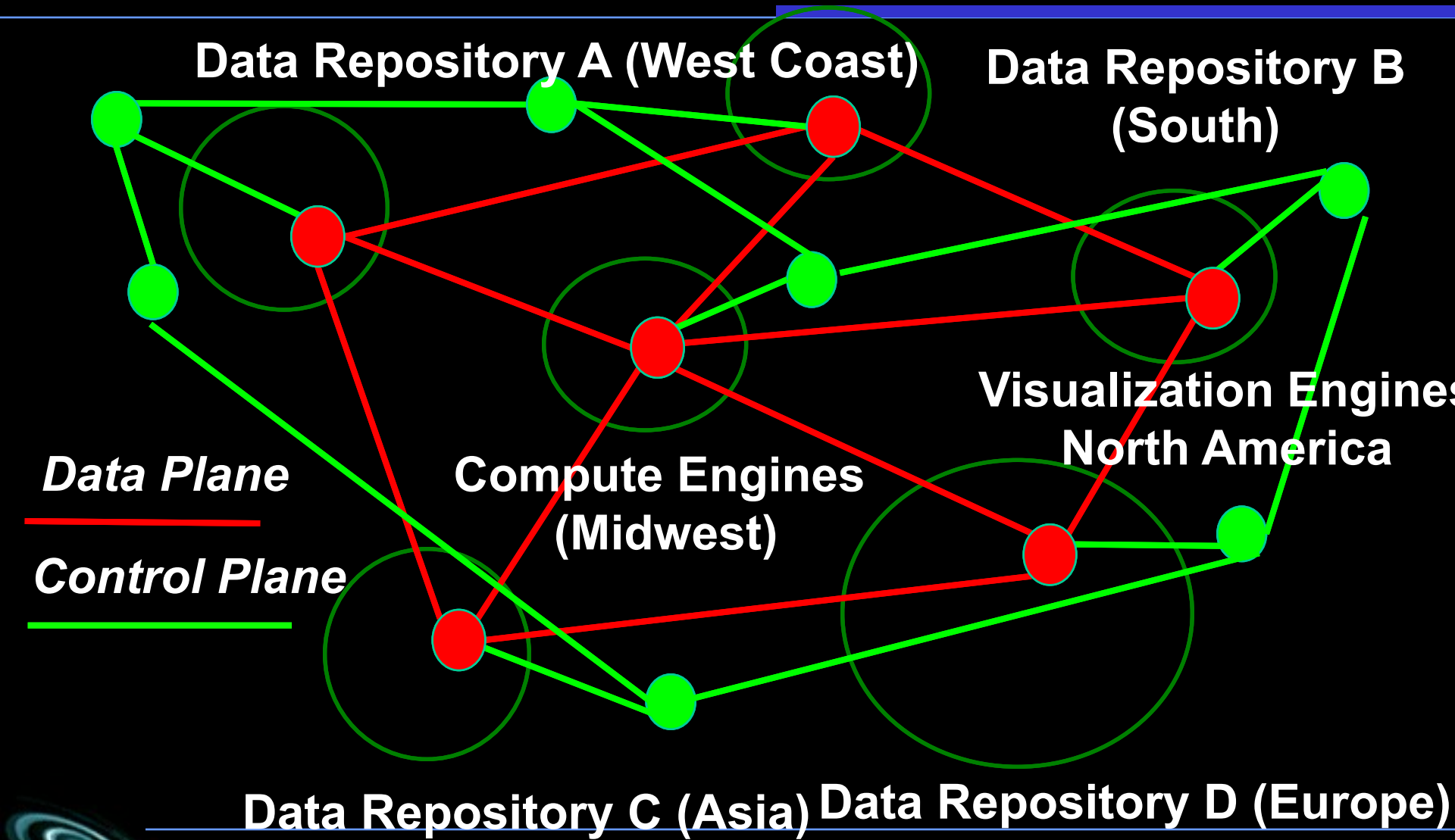
NCI Genomic Data Commons



- Harmonization and storage for the Nations Cancer Genomic Data GDC 1.6PB of cancer genomic data and associated clinical data.
- Precision Medicine Enabled By Precision Networking

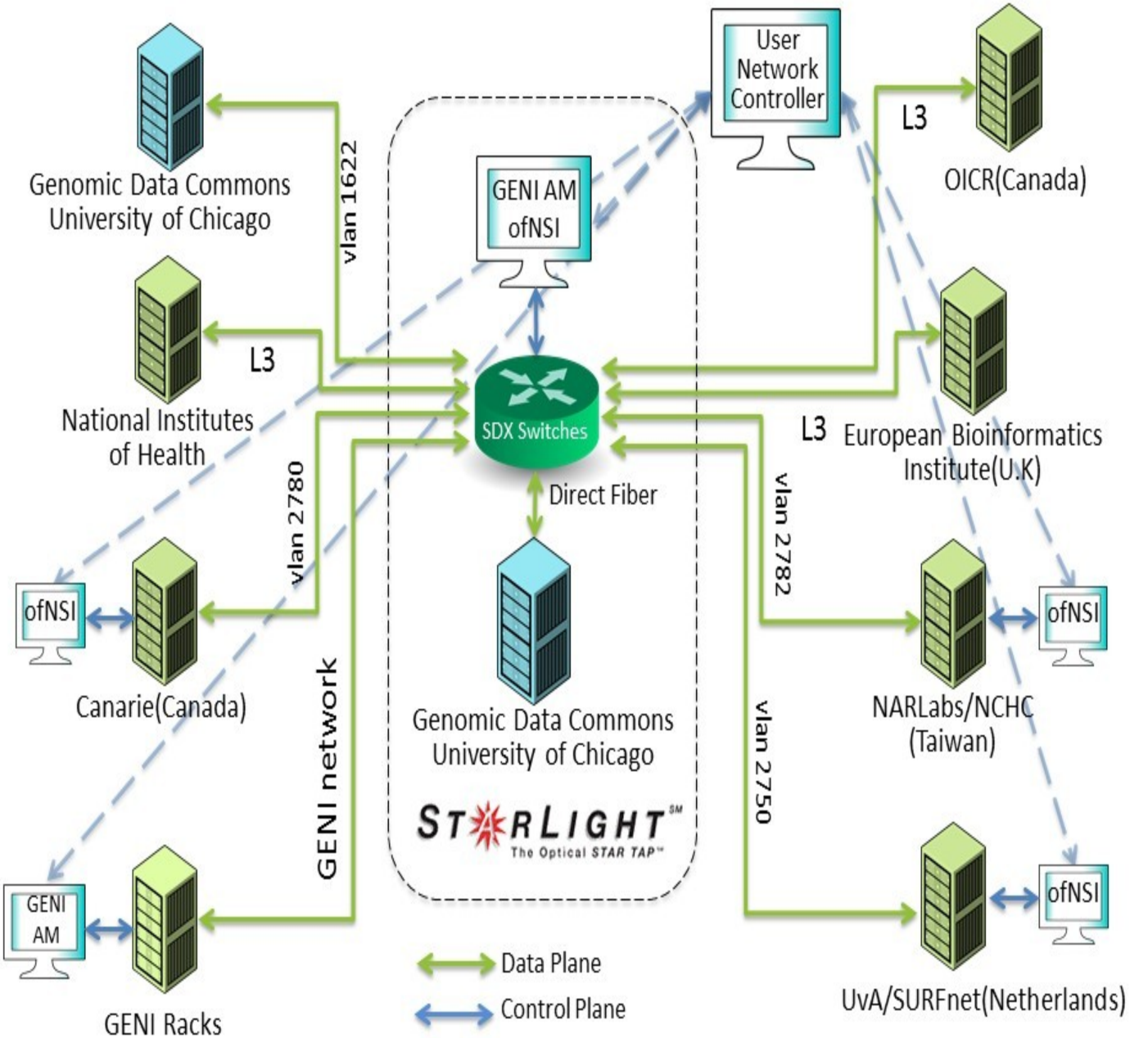
STARLIGHT℠

Biomedical Data Commons:
Flow Orchestration: Control Plane + Data Plane

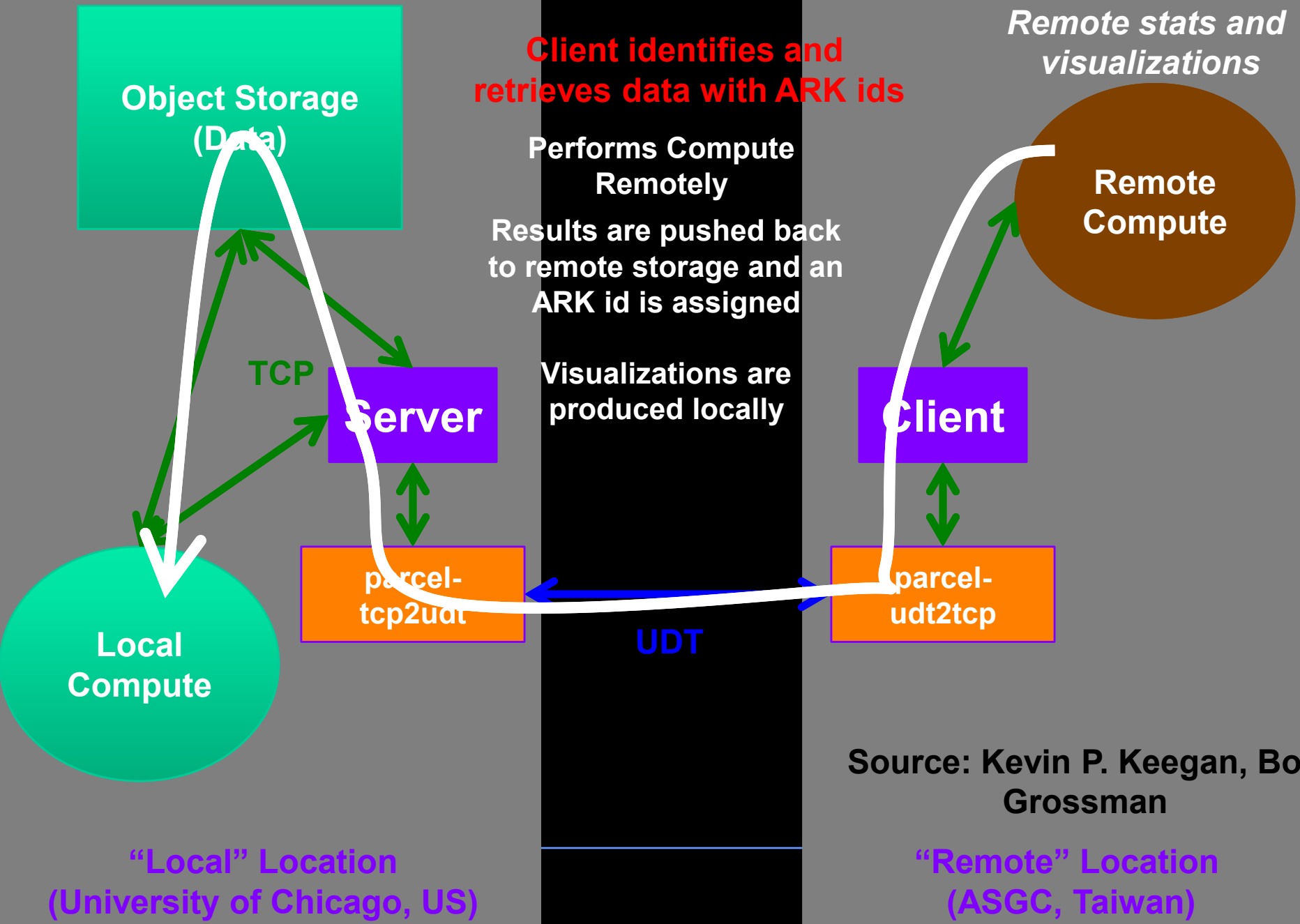


STARLIGHT℠

GEC22 Bioinformatics SDXs Demo Network

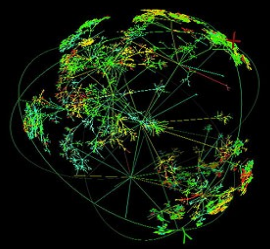


Parcel Based Collaboration



IRNC: RXP: StarLight SDX A Software Defined Networking Exchange for Global Science Research and Education

Joe Mambretti, Director, (j-mambretti@northwestern.edu)
International Center for Advanced Internet Research (www.icair.org)
Northwestern University
Director, Metropolitan Research and Education Network (www.mren.org)
Co-Director, StarLight (www.startap.net/starlight)
PI IRNC: RXP: StarLight SDX
Co-PI Tom DeFanti, Research Scientist, (tdefanti@soe.ucsd.edu)
California Institute for Telecommunications and Information Technology (Calit2),
University of California, San Diego
Co-Director, StarLight
Co-PI Maxine Brown, Director, (maxine@uic.edu)
Electronic Visualization Laboratory, University of Illinois at Chicago
Co-Director, StarLight
Jim Chen, Associate Director, International Center for Advanced Internet Research, Northwestern University



National Science Foundation
International Research Network Connections Program
Workshop
Chicago, Illinois
May 15, 2015

Another SDX Opportunity! An Experimental Testbed For Computer Science Research

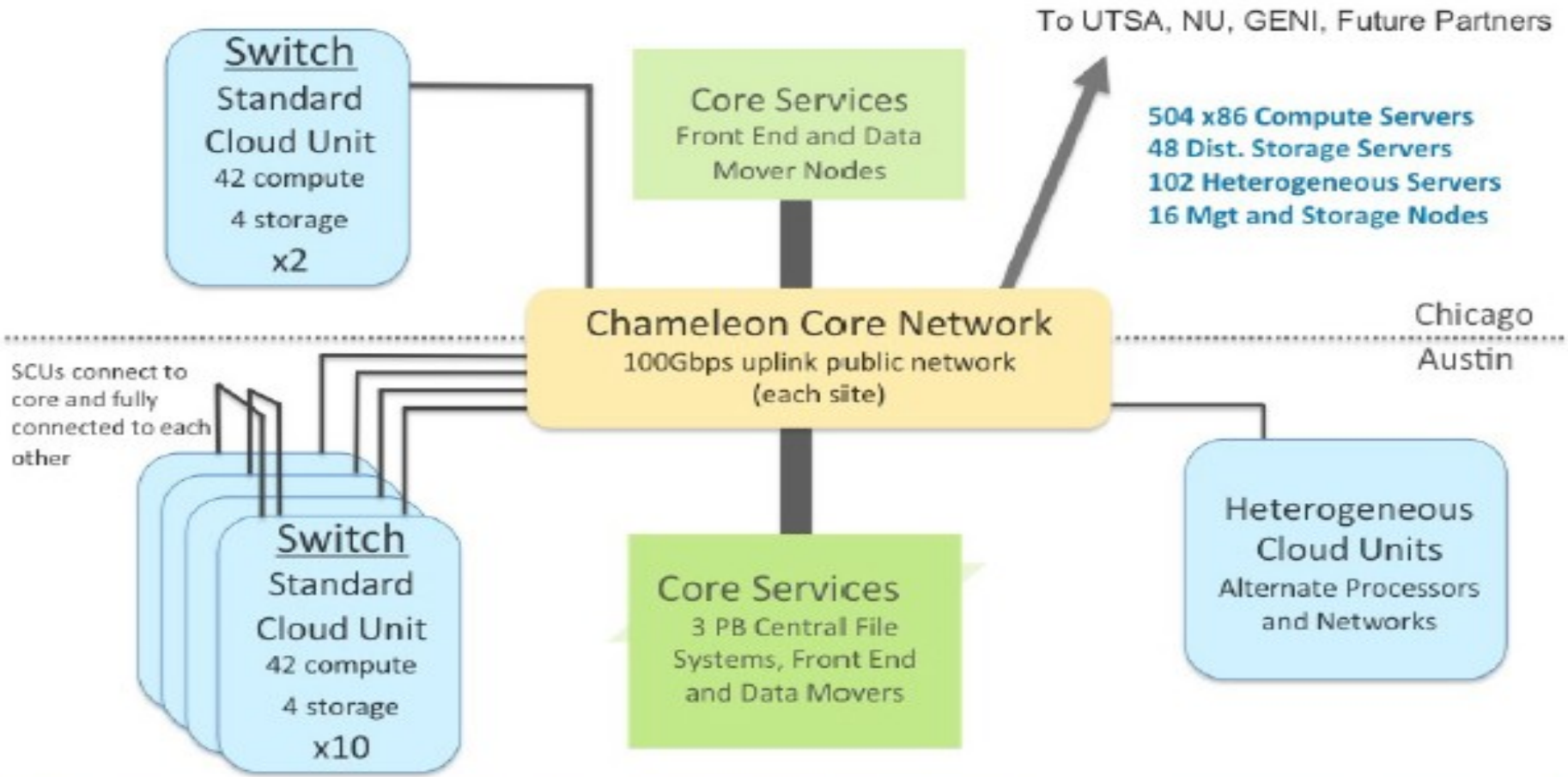
www.chameleoncloud.org

CHAMELEON:
A LARGE-SCALE, RECONFIGURABLE EXPERIMENTAL ENVIRONMENT FOR CLOUD RESEARCH

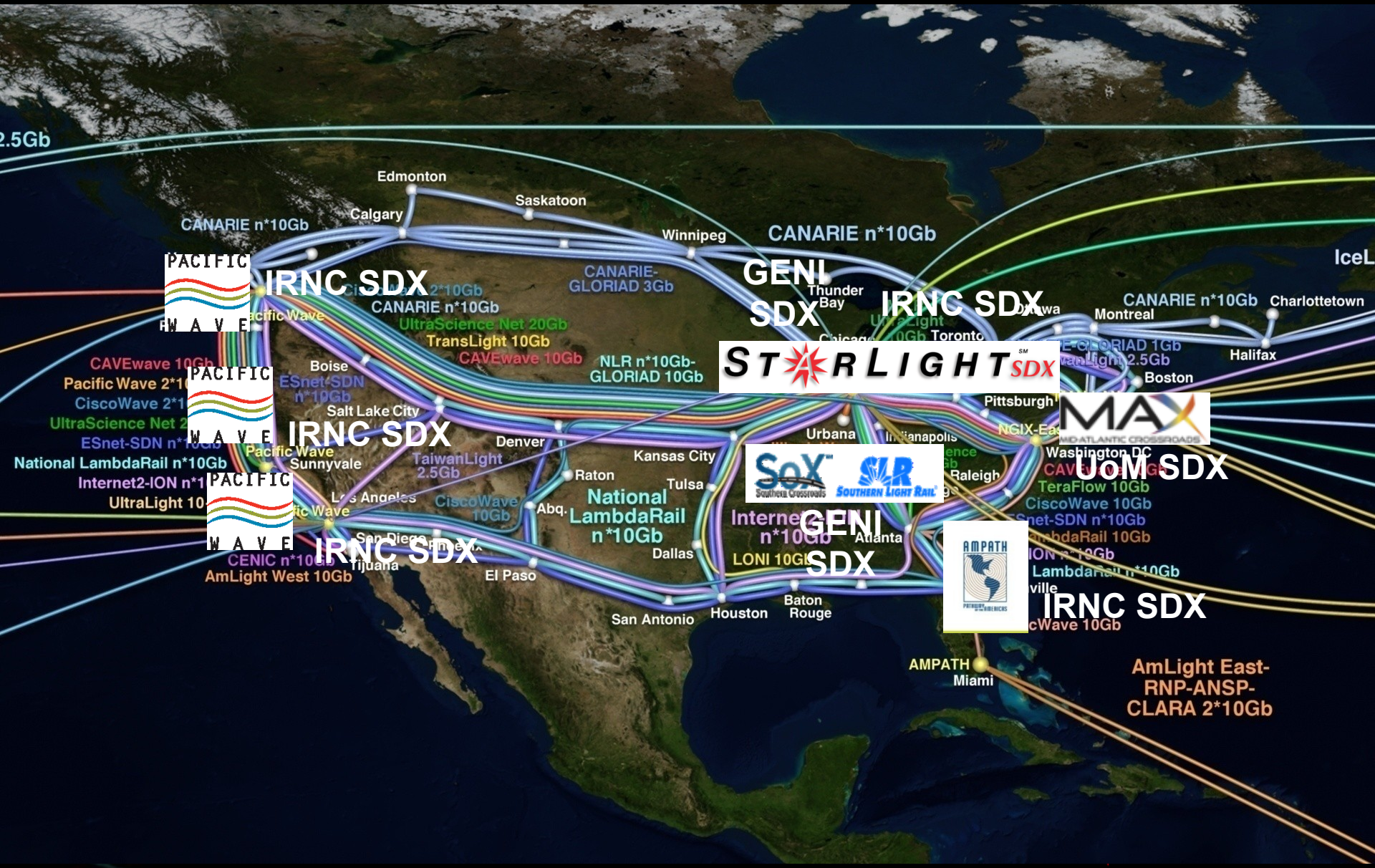
Principal Investigator: Kate Keahey

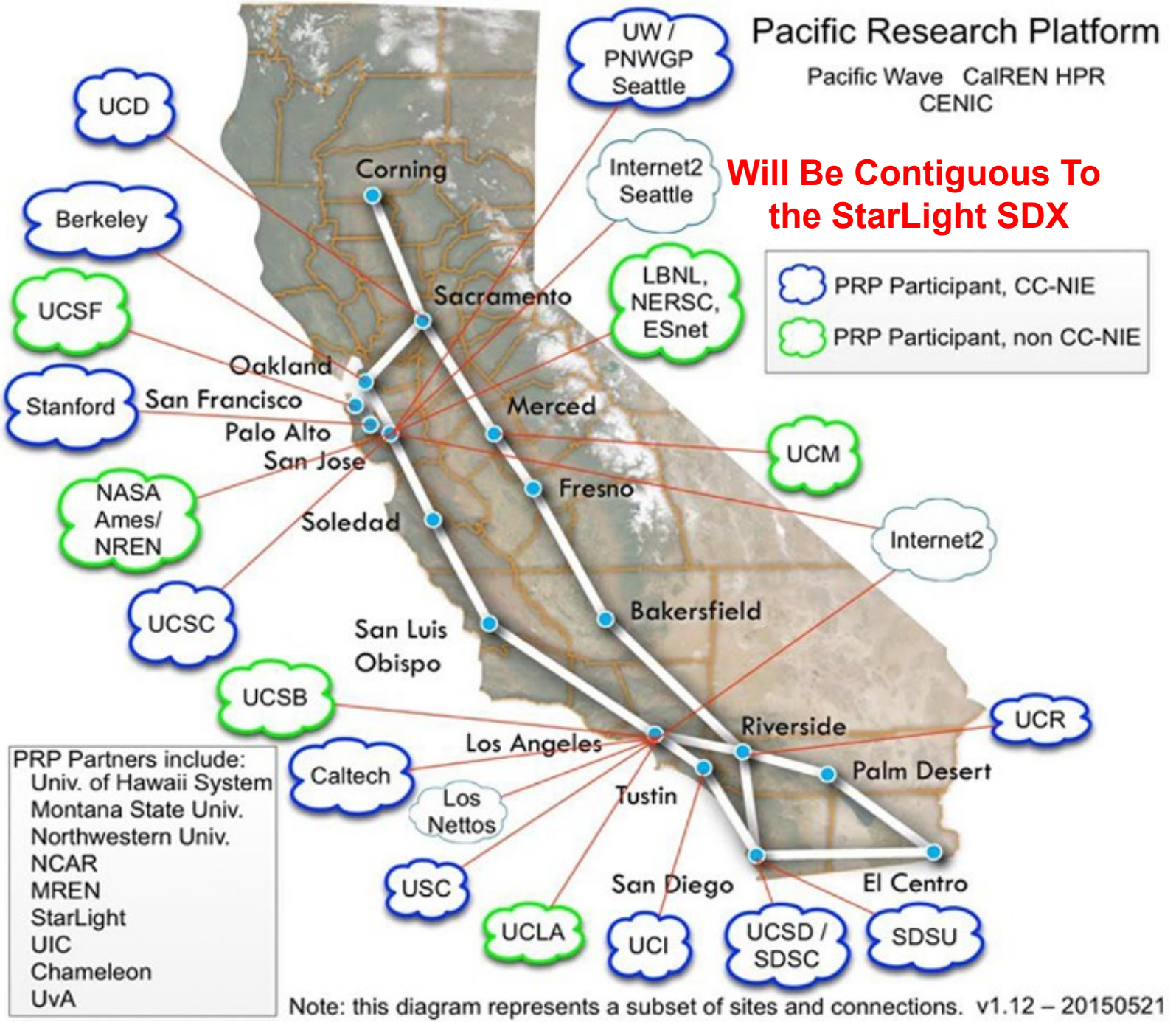
Co-PIs: J. Mambretti, D.K. Panda, P. Rad, W. Smith, D. Stanzione

CHAMELEON ENVIRONMENT



Planned US SDX Interoperable Fabric




Pacific Research Platform
Pacific Wave CalREN HPR
CENIC
Will Be Contiguous To
the StarLight SDX
PRP Participant, CC-NIE
PRP Participant, non CC-NIE
UW / PNWGP Seattle
Internet2 Seattle
LBNL, NERSC, ESnet
UCD
Berkeley
UCSF
Stanford
NASA Ames/ NREN
UCSC
UCSB
Caltech
Los Nettos
USC
UCLA
UCI
UCSD / SDSC
SDSU
UCM
Internet2
UCR
Corning
Sacramento
Oakland
San Francisco
Palo Alto
San Jose
Soledad
Merced
Fresno
Bakersfield
San Luis Obispo
Los Angeles
Riverside
Palm Desert
Tustin
San Diego
El Centro
PRP Partners include:
Univ. of Hawaii System
Montana State Univ.
Northwestern Univ.
NCAR
MREN
StarLight
UIC
Chameleon
UvA
Note: this diagram represents a subset of sites and connections. v1.12 – 20150521

Next Step: Global Research Platform Building on CENIC/Pacific Wave and GLIF

Current International GRP Partners

Global Research Platform

- A Emerging International Fabric
- A Specialized Globally Distributed Platform For Science Discovery and Innovation
- Based On State-Of-the-Art-Clouds
- Interconnected With Computational Grids, Supercomputing Centers, Specialized Instruments, et al
- Also, Based On World-Wide 100 Gbps Networks
- Leveraging Advanced Architectural Concepts, e.g., SDN/SDX/SDI – Science DMZs
- Ref: Demonstrations @ SC15, Austin Texas November 2015
- ***New=> Global Research Platform 100 Gbps Network (GRPnet) On Private Optical Fiber Between PacificWave and StarLight via the PNWGP***

www.startap.net/starlight